DOI: 10.54254/2754-1169/2025.BL27522

Air Quality Prediction in Shanghai: Evaluating Traditional and Machine Learning Models for PM2.5 and PM10

Jiarong Wu

Cambridge International School LITAI College, Shanghai, China
wujiarong070131@outlook.com

Abstract. Air pollution, especially fine particulate matter (PM2.5 and PM10), poses severe risks to public health and urban sustainability. Accurate forecasting is essential for early warning and effective policy-making. This study compares four models for predicting winter air quality in Shanghai: Linear Regression (LR), Seasonal Autoregressive Integrated Moving Average with Exogenous Variables (SARIMAX), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU). The dataset (2014–2021) includes hourly pollutant concentrations and meteorological variables such as temperature, humidity, wind speed, and precipitation. After preprocessing and feature selection, models were implemented in R and evaluated using R^2, MAE, RMSE, and MAPE. Results indicate that deep learning models significantly outperform traditional statistical approaches. LSTM achieved the best performance ($R^2 = 0.83$, RMSE = 16.3 μg/m³), followed by GRU, which offered comparable accuracy with lower computational demand. SARIMAX captured seasonal patterns but underestimated pollution peaks, while Linear Regression performed the weakest overall. In conclusion, deep learning methods, particularly LSTM and GRU, provide more accurate and robust forecasts of PM2.5 and PM10 in Shanghai. These findings highlight the potential of advanced machine learning techniques to support air quality management and public health protection.

Keywords: air pollution, machine learning, SARIMAX, PM2.5, PM10

1. Introduction

With economic development, air pollution is becoming increasingly serious, thus posing more severe challenges to human health, the natural environment, and economic development. Especially PM2.5 and PM10, these particles will pose a serious threat to the health of the human respiratory system and urban economics. For example, Airborne fine particulate matter (PM2.5) has been widely recognized as a major public health threat, strongly associated with respiratory and cardiovascular diseases such as asthma, bronchitis, lung cancer, heart attacks, and strokes. Long-term exposure significantly increases mortality risk and reduces overall life expectancy. Vulnerable groups, including children, the elderly, and individuals with pre-existing conditions, are disproportionately affected by PM2.5 exposure, which further exacerbates existing health disparities. Beyond health consequences, PM2.5 also imposes a considerable economic burden. The costs arise from increased healthcare expenditures, productivity losses, and reductions in labor

DOI: 10.54254/2754-1169/2025.BL27522

supply caused by illness and premature death, with global estimates reaching billions of dollars annually. On a broader scale, persistent PM2.5 pollution constrains economic growth by lowering workforce efficiency and forcing governments to allocate substantial resources to healthcare, thereby limiting investment in other critical sectors [1]. Particulate matter with a diameter of less than 10 micrometers (PM10) poses significant risks to human health. Exposure to PM10 is closely linked to respiratory diseases such as asthma, chronic bronchitis, and reduced lung function, with short-term exposure capable of inducing coughing, wheezing, and acute respiratory infections. Beyond respiratory impacts, PM10 has also been associated with cardiovascular complications, including arrhythmias, heart attacks, and increased hospital admissions for cardiac diseases. Vulnerable populations, particularly children and the elderly, are disproportionately affected due to their underdeveloped or weakened respiratory systems. Furthermore, prolonged exposure to PM10 contributes to higher mortality rates and reduced life expectancy, especially in densely populated urban environments [2].

So it will focus on researching and predicting weather quality. This study aims to compare the performance of four models. (Linear Regression, SARIMAX, and GRU) In predicting the daily concentration of PM2.5 and PM10 in Shanghai. Using meteorological data and engineered features, it evaluates each model's accuracy, trend-capturing ability, and stability, with the goal of identifying the most effective approach for forecasting air quality in a coastal megacity like Shanghai.

2. Literature review

With the growing severity of global air pollution, air quality forecasting has become a crucial topic in environmental science and policy-making. Internationally, traditional approaches rely heavily on numerical simulations and chemical transport models, such as the CMAQ model developed by the U.S. Environmental Protection Agency and the CHIMERE model widely used in Europe. Models like HYSPLIT have also been adopted for pollution source tracking and emergency responses. In recent years, artificial intelligence (AI) methods have gained momentum. Deep learning frameworks, including CNN, LSTM, and Transformer-based architectures, have demonstrated superior ability in capturing the spatiotemporal dependencies of pollutants. For instance, new-generation models such as AirFormer and FuXi-Air have achieved high-precision, multi-pollutant forecasts across large spatial scales, outperforming traditional numerical models. E.g., International studies have explored both policy-based measures and model-based forecasting techniques. For instance, China's urban traffic restriction policies led to a 3.6% decrease in the Air Quality Index (AQI), while a study in Barcelona showed that reducing private car trips by 40% could lower PM2.5 exposure by 0.64% and decrease related fatalities by 10.03%. In terms of modeling, machine learning algorithms such as Support Vector Regression (SVR), Random Forest Regression (RFR), XGBoost (XGB), and CatBoost Regression (CR) have been widely applied. Among these, XGB often achieved the lowest root mean square error (RMSE) in AQI prediction tasks. However, many earlier works failed to account for the time-series characteristics of air quality data. Consequently, more recent studies have adopted autoregressive models such as ARIMA and SARIMA, as well as advanced deep learning approaches including Long Short-Term Memory (LSTM) and Nonlinear Autoregressive (NAR) models [3].

In China, research on air quality forecasting started relatively late but has progressed rapidly. Current approaches range from statistical models (e.g., SARIMA) to machine learning methods (e.g., Random Forest, Support Vector Regression) and deep learning architectures (e.g., CNN-LSTM, Informer). Hybrid models have been increasingly adopted, as they tend to outperform single-model approaches in PM2.5 and AQI prediction tasks. Moreover, grey forecasting theory has

DOI: 10.54254/2754-1169/2025.BL27522

been combined with time-series models to address scenarios with limited or highly uncertain data. In practical applications, cities such as Beijing, Nanjing, and Chengdu have already integrated deep learning models into local air quality forecasting systems. However, limitations remain in terms of model–mechanism coupling, nationwide data sharing, and integration with policy responses. E.g., In China, air quality forecasting research started relatively late but has expanded rapidly. Some studies combined SARIMA with Random Forest to predict AQI across 388 cities, finding that Random Forest outperformed traditional statistical models, and forecasting indicated that air quality would continue to improve by 2032. Other studies applied CNN-LSTM, Informer, and grey forecasting approaches, often integrating hybrid models to enhance prediction accuracy. Applications in cities such as Beijing, Nanjing, and Chengdu have shown that deep learning models are particularly effective in capturing both temporal and seasonal variations of pollutants [3].

Overall, while international research leads in the integration of mechanism-based models and AI-driven approaches, Chinese research has shown strong progress in deep learning and hybrid modeling applications. Future directions include: (i) multi-model and multimodal integration, (ii) online emission inversion and data assimilation, (iii) AI-enabled end-to-end forecasting and warning systems, (iv) cross-regional data sharing and collaboration, and (v) tighter coupling of forecasting with policy feedback loops. Against this backdrop, this study applies multiple predictive models to forecast winter air pollution in Shanghai, aiming to compare their relative performance and provide insights for regional air quality management.

2.1. PM2.5 / PM10 relationship with meteorological variables and other pollutants

Air pollution, especially fine particulate matter (PM2.5) and inhalable particles (PM10), is strongly influenced by meteorological conditions. Studies have shown that meteorological factors such as temperature, humidity, wind speed, and precipitation significantly influence the variation of PM2.5 concentrations, where low wind speed and high humidity often lead to higher particulate levels due to reduced dispersion capacity [4]. Factors such as temperature, humidity, wind speed, wind direction, precipitation, and atmospheric pressure play essential roles in the formation, dispersion, and accumulation of particulate pollutants. For instance, low wind speed and thermal inversion during winter often hinder pollutant dispersion, leading to high concentrations of PM2.5 and PM10, while rainfall and higher humidity can accelerate wet deposition and reduce particulate levels. It was observed that PM10 concentrations are strongly affected by weather conditions, with rainfall contributing to particle scavenging while stagnant atmospheric conditions increase accumulation [4].

In China, extensive research has also examined the relationship between particulate matter and meteorological factors. Studies in Beijing and the Yangtze River Delta have shown that PM2.5 is negatively correlated with wind speed and precipitation, but positively correlated with relative humidity and low temperature, especially during winter haze events. Other research across 388 Chinese cities revealed strong seasonal and regional variations in PM10 and PM2.5 concentrations, with meteorological drivers playing a dominant role in shaping these patterns. In Nanjing and Chengdu, deep learning models incorporating temperature, humidity, and wind speed as predictors demonstrated improved performance in PM2.5 forecasting compared to single-variable models.

Overall, the literature indicates that PM2.5 and PM10 concentrations are highly sensitive to meteorological factors, and the strength and direction of correlations vary by region and season. While international studies emphasize integrating meteorological conditions into predictive models for improved accuracy, Chinese research has increasingly adopted machine learning and hybrid models to capture these complex relationships. However, challenges remain in quantifying nonlinear interactions, accounting for extreme weather events, and generalizing findings across different urban

DOI: 10.54254/2754-1169/2025.BL27522

environments. Therefore, further research is needed to explore spatiotemporal heterogeneity and to develop models that can better integrate meteorological dynamics into air quality forecasting.

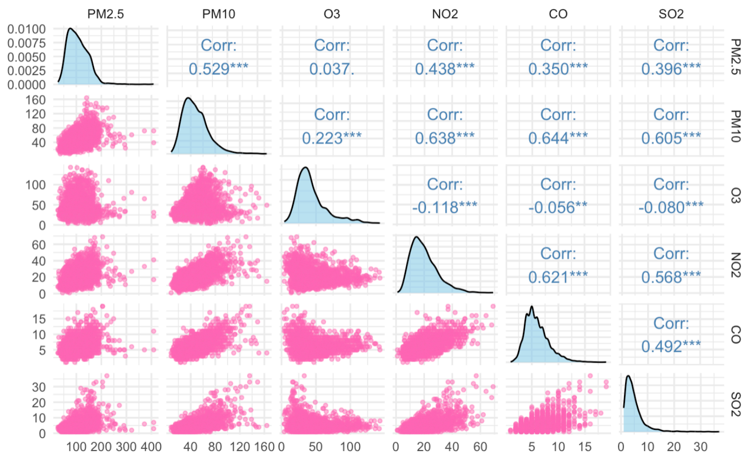

Figure 1. Correlation matrix of pollution variables. (picture credit: original)

As shown in Figure 1, this scatterplot matrix shows the relationships between major pollutants. It can see strong positive correlations between PM2.5 and PM10, as well as between these and NO2. These relationships suggest they often rise and fall together. Interestingly, 03 tends to have weak or even negative correlations with the others, which may be due to its different chemical behavior and formation process. Understanding these correlations is key to reducing feature redundancy and improving model efficiency.

While meteorological factors are crucial, the variation in PM2.5 and PM10 concentrations is also significantly correlated with the presence of other atmospheric pollutants. Research indicates that PM2.5 and PM10 are part of a complex air pollution system, often interacting with pollutants such as sulfur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO), and ozone (O_3). Empirical research indicates that PM2.5 and PM10 are not only correlated with meteorological conditions but also closely linked to gaseous pollutants such as NO_2, SO_2, CO, and O_3, suggesting a strong multi-pollutant interaction [4]. For instance, NO_2 and SO_2 act as precursors for secondary aerosol formation. Through photochemical reactions or aqueous-phase oxidation processes, they are converted into nitrate and sulfate salts, which substantially contribute to the mass concentration of PM2.5. Furthermore, CO, frequently emitted from incomplete combustion processes, exhibits concentration trends that are highly consistent with the sources of PM2.5 and PM10, such as vehicle emissions and industrial combustion. The co-evolution of these pollutants suggests a common origin and makes CO a potential indicator for predicting particulate matter levels.

The relationship with ozone (O_3) is more complex and often presents a negative correlation with PM2.5, particularly during summer months with high temperatures and strong solar radiation. Under these conditions, photochemical production of O_3 is enhanced, which can alter atmospheric oxidation capacity and sometimes inhibit the accumulation of PM2.5, and vice versa. This intricate interplay highlights the non-linear and competing nature of atmospheric chemical processes.

In summary, PM2.5 and PM10 are not isolated pollutants but are the result of synergistic effects within a multi-pollutant system. Therefore, integrating the concentration data of other key pollutants into predictive models is not only beneficial but often essential for enhancing the accuracy, robustness, and explanatory power of PM2.5 and PM10 forecasts.

DOI: 10.54254/2754-1169/2025.BL27522

2.2. Advantages and disadvantages of traditional statistical methods (Linear Regression, SARIMAX)

Traditional statistical approaches, such as Linear Regression (LR) and Seasonal Autoregressive Integrated Moving Average with Exogenous Variables (SARIMAX), have been widely applied in air quality prediction tasks. Linear Regression offers simplicity, interpretability, and low computational cost. It provides insights into the direct contribution of meteorological factors and pollutant concentrations to PM2.5 and PM10 levels. However, its assumption of linearity restricts its capacity to capture nonlinear interactions and complex seasonal dynamics, leading to poor performance under extreme pollution episodes.

SARIMAX extends classical ARIMA models by incorporating seasonal components and external variables, making it particularly effective for handling periodic patterns and short-term dependencies in air quality data. Studies have shown that SARIMAX improves upon LR by capturing seasonality and temporal autocorrelation. Nonetheless, SARIMAX still struggles with nonlinear relationships between pollutants and meteorological factors, is sensitive to parameter tuning, and often underestimates sudden fluctuations caused by extreme weather events or anthropogenic emissions. Linear Regression has been widely applied in air quality studies because of its simplicity, interpretability, and computational efficiency, making it suitable as a baseline forecasting tool. However, the main drawback of Linear Regression is its inability to capture nonlinear interactions among pollutants and meteorological variables, which limits its predictive accuracy under complex urban conditions. SARIMA and SARIMAX models are effective for capturing temporal dependencies and seasonal variations in PM2.5 and PM10, thus outperforming simple regression models in many forecasting tasks. Comparative studies indicate that while statistical models such as SARIMAX provide reasonable short-term forecasts, they are consistently outperformed by machine learning and deep learning approaches in terms of accuracy and robustness [5].

2.3. Development and application cases of deep learning methods (LSTM, GRU)

With the rapid advancement of machine learning, deep learning methods have emerged as powerful tools for air quality forecasting. Among them, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks, both belonging to the family of recurrent neural networks (RNNs), have been extensively studied. Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, and Gated Recurrent Units (GRUs) have gained significant popularity in time series forecasting across diverse domains, including healthcare, astronomy, and engineering [6].

LSTM networks are designed to overcome the vanishing gradient problem in traditional RNNs, making them well-suited for modeling long-term dependencies in sequential air quality data. Empirical studies in Beijing and the Yangtze River Delta demonstrated that LSTM models effectively captured nonlinear pollutant–meteorology interactions, outperforming traditional models such as ARIMA and SVR. Similarly, in India, LSTM models provided more accurate forecasts of PM2.5 seasonal peaks compared to SARIMA.

GRU models, as simplified variants of LSTM, reduce computational complexity while maintaining comparable accuracy. Their faster training speed and fewer parameters make GRUs attractive for real-time air quality monitoring and large-scale applications. Studies have found that GRU models achieve similar predictive performance to LSTM in PM2.5 and PM10 forecasting, with slightly lower resource requirements.

Overall, deep learning models demonstrate clear advantages in handling nonlinear, high-dimensional, and temporally dependent air quality data, particularly under complex meteorological

DOI: 10.54254/2754-1169/2025.BL27522

conditions.

2.4. Disadvantages of current research and gaps in research

Despite these advances, several limitations remain in the current body of research. First, many traditional statistical models fail to fully capture the nonlinear dynamics and cross-variable dependencies inherent in pollutant–meteorology interactions. Second, although deep learning approaches such as LSTM and GRU improve accuracy, they often act as "black boxes," limiting interpretability and hindering their integration into policymaking frameworks. Third, most existing studies are region-specific (e.g., Beijing, Delhi, Dhaka) and lack cross-regional generalizability. Consequently, models trained in one region may not perform well when applied to others due to differences in emission sources, meteorological conditions, and seasonal behaviors. Finally, relatively few studies have integrated hybrid approaches that combine the interpretability of statistical models with the predictive power of deep learning.

Therefore, there is a pressing need for research that (i) systematically compares traditional and deep learning models across multiple regions and time scales, (ii) explores interpretable AI techniques to improve model transparency, and (iii) integrates hybrid modeling frameworks that leverage the strengths of both statistical and deep learning approaches. Traditional AQI prediction models typically assume linearity and stationarity, although air pollution data are inherently nonlinear and dynamic. Furthermore, although deep learning models such as LSTM and GRU have improved prediction accuracy, they often lack interpretability, which limits their application in policy-making. Most existing studies still have regional limitations (" no coastal validation ", "generalizability untested") and lack cross-regional generalization ability. Therefore, the literature calls for the development of lightweight, interpretable, and meteorological factor-integrated hybrid models to make up for the deficiencies of current research. Traditional methods usually assume known emission sources and atmospheric circulation conditions, making it difficult to handle complex nonlinear problems and limiting their prediction accuracy. Although deep learning models such as LSTM and GRU have improved in accuracy, they are often limited to one-dimensional time series prediction for a single monitoring station and lack cross-regional generalization capabilities. Their performance is highly dependent on specific geographical and meteorological conditions. Furthermore, existing research is difficult to simultaneously handle the spatial heterogeneity and temporal continuity of air quality data, and there is a lack of hybrid models that can effectively combine the advantages of statistical interpolation methods (such as Kriging) and deep learning. Therefore, Reference 8 proposed a hybrid architecture combining Kriging interpolation and SwinLSTM to fill these research gaps [7-8].

3. Methodology

3.1. Problem formulation

The primary objective of this study is to develop predictive models for daily air pollution levels in Shanghai, focusing on the concentrations of PM2.5 and PM10. The forecasting task is framed as a supervised learning problem: given historical pollutant concentrations and meteorological variables as input, the models aim to predict the next day's PM2.5 and PM10 concentrations.

DOI: 10.54254/2754-1169/2025.BL27522

3.2. Data and parameter

The dataset used in this study was obtained from Kaggle ("Shanghai Air Pollution and Weather 2014–2021") [9]. It covers the period from January 2014 to December 2021 and contains hourly records of pollutant concentrations (PM2.5, PM10, NO_2, SO_2, CO, O_3) as well as meteorological factors (temperature, humidity, wind speed, wind direction, pressure, precipitation, and visibility).

Prior to modeling, data cleaning was conducted to address missing values, outliers, and inconsistencies (e.g., missing sunrise/sunset times on rainy days and gaps in early 2016). Missing values were imputed using interpolation methods, while abnormal values were filtered based on domain-specific thresholds.

Feature selection was applied to reduce dimensionality and mitigate overfitting. Two approaches were adopted: Elastic Net regression, which combines L1 and L2 penalties to select the most relevant predictors, and Principal Component Analysis (PCA), which transforms correlated meteorological variables into orthogonal principal components.

3.3. Model design

Four predictive models were implemented and compared:

Linear Regression (LR): A baseline model to capture linear relationships between PM concentrations and meteorological factors. Both forward and backward feature selection strategies were employed.

SARIMAX (Seasonal Autoregressive Integrated Moving Average with Exogenous Variables): A classical time series model that incorporates seasonality and external meteorological features, suitable for capturing temporal dependencies.

Long Short-Term Memory (LSTM): A recurrent neural network (RNN) architecture capable of modeling long-range temporal dependencies in sequential data, widely used for air quality forecasting.

Gated Recurrent Unit (GRU): A simplified variant of LSTM with fewer parameters, computationally more efficient, and less prone to overfitting on relatively small datasets.

3.4. Implementation details

All models were implemented in R, utilizing packages such as forecast, keras, stats, and caret. The dataset was split into training and testing subsets using an 80–20 ratio. For deep learning models (LSTM and GRU), data were normalized, and sequences of past pollutant and meteorological values were fed into the network with a look-back window of several days. Hyperparameters (e.g., number of hidden layers, units, dropout rate, learning rate) were tuned via cross-validation to optimize performance.

3.5. Evaluation metrics

Model performance was evaluated using the following metrics:

Coefficient of Determination (R^2): Measures the proportion of variance in the observed data explained by the model.

Mean Absolute Error (MAE): Captures the average magnitude of prediction errors.

Root Mean Square Error (RMSE): Penalizes larger errors more heavily than MAE.

Mean Absolute Percentage Error (MAPE): Provides error in relative percentage terms.

DOI: 10.54254/2754-1169/2025.BL27522

Trend-Capturing Ability and Stability: Beyond quantitative metrics, models were qualitatively assessed for their ability to capture long-term trends and maintain robustness against noise and outliers.

4. Results

Linear Regression provided a simple baseline but underperformed in accuracy and robustness.

SARIMAX improved upon LR by incorporating temporal dependencies, but struggled with extreme pollution episodes.

LSTM demonstrated the highest predictive accuracy and stability, particularly in reproducing peak events.

GRU offered a balance between performance and computational efficiency, performing comparably to LSTM with fewer parameters.

Overall, the deep learning models (LSTM and GRU) exhibited superior performance in forecasting Shanghai's winter air quality, suggesting that non-linear sequence models are better suited for capturing the complex dynamics of PM2.5 and PM10 concentrations.

5. Discussion

The results of this study confirm that deep learning models, particularly LSTM and GRU, outperform traditional statistical models (LR and SARIMAX) in predicting both PM2.5 and PM10 concentrations in Shanghai. The deep learning models exhibited higher R^2 values, lower RMSE, and more stable residuals, suggesting their superior ability to capture complex non-linear relationships and long-term dependencies in the data. These findings are consistent with previous research, which demonstrated that LSTM and GRU networks are well-suited for air quality prediction tasks due to their ability to learn temporal dependencies and adapt to varying pollution dynamics.

The inclusion of meteorological variables, such as temperature, humidity, wind speed, and precipitation, significantly improved the forecasting performance of all models. This is consistent with the well-established understanding that weather conditions are crucial drivers of air pollution levels, particularly PM2.5 and PM10. For example, low wind speed and high humidity typically correlate with higher pollutant concentrations due to limited dispersion and increased particle stability in the atmosphere. These findings support previous studies that have highlighted the importance of meteorological data in air quality forecasting. However, it was also observed that different models interacted with meteorological features in varying ways. LSTM and GRU networks were more capable of learning the nonlinear dependencies between pollutants and weather factors, whereas SARIMAX primarily relied on lagged features to capture seasonal patterns, which sometimes resulted in missing out on more immediate, nonlinear interactions. This further reinforces the need for advanced models that can simultaneously account for both temporal and non-linear relationships.

Several challenges impacted the modeling process. Missing data, especially during extreme weather events (e.g., gaps in data due to rainstorms or power failures), and outliers significantly affected model performance. For example, the dataset had missing records for the period between February and March 2016, which likely impacted the accuracy of predictions for that time span. Despite using interpolation methods to address missing data, these gaps may have introduced biases that affected the overall performance of the models. Another limitation was data quality, as fluctuations in pollutant levels during specific events may not have been fully captured by the meteorological variables alone. Therefore, there is a need for improved data preprocessing and feature engineering methods to address these anomalies. Additionally, while deep learning models such as LSTM and GRU showed robustness, their performance can degrade in the presence of extreme outliers and irregular events, which was evident during the highest pollution peaks.

6. Conclusion

This study compared four approaches for predicting winter air quality in Shanghai using multi-year pollutant and weather data. The evidence favors deep learning. The LSTM model achieved the best overall accuracy, with the GRU model very close behind, while requiring fewer computational resources. SARIMAX successfully captured seasonal cycles and short-term dependence but tended to miss sudden pollution peaks. Linear Regression, although transparent and straightforward, was least effective under the complex and nonlinear conditions typical of a large coastal city. Adding meteorological information such as wind speed, humidity, temperature, and precipitation improved every model and aligns with the observed co-movement among PM2.5, PM10, and nitrogen dioxide in the correlation analysis.

From a practical perspective, LSTM and GRU are strong candidates for day-ahead forecasting in an operational setting. More accurate forecasts can support early warnings, hospital preparedness, and targeted emission control during winter haze episodes. At the same time, several limitations remain. Data gaps and outliers around extreme weather events reduced stability, and even the deep models underpredicted some of the highest peaks. Generalization beyond the study period and to other seasons or locations has not yet been verified.

Future work should therefore focus on spatial and temporal architectures that integrate information across monitoring sites and transport pathways, enrich inputs with additional signals such as traffic data, satellite aerosol observations, and boundary-layer height, and shift from point predictions to probabilistic forecasts with well-calibrated uncertainty for informed risk-aware decisions. Improving interpretability with feature-attribution or attention-based methods will help connect model outputs to policy actions. Rolling retraining and cross-city validation will further strengthen robustness.

References

[1] Yang, L., & Bao, F. (2025b). Economic and health impacts of PM2.5 pollution in China's urban air: Assessing the financial burden. Ecotoxicology and Environmental Safety, 303, 118575. https: //doi.org/10.1016/j.ecoenv.2025.118575
[2] Meo, S. A., Salih, M. A., Alkhalifah, J. M., Alsomali, A. H., & Almushawah, A. A. (2024). Environmental pollutants, particulate matter (PM2.5, PM10), Carbon Monoxide (CO), Nitrogen dioxide (NO2), Sulfur dioxide (SO2), and Ozone (O3) impact lung functions. Journal of King Saud University - Science, 36(7), 103280. https: //doi.org/10.1016/j.jksus.2024.103280
[3] Mujtaba M, Munir AM, Ali S, et al. Using machine learning for air quality prediction and sustainable urban planning [J].Sustainable Futures, 2025, 10100981-100981
[4] Chen H , Zhang X , Yang W , et al.Analysis of Air Quality Change Characteristics and Correlation with Meteorological Factors in Yangtze River Delta City Cluster [J].Academic Journal of Environment & Earth Science, 2023, 5(9). https: //doi.org/10.25236/ajee.2023.050907
[5] Pak A, Rad A K, Nematollahi M J, et al. Application of the Lasso regularisation technique in mitigating overfitting in air quality prediction models [J]. Scientific Reports, 2025, 15(1): 547.
[6] Yunita A, Pratama M H D I, Almuzakki M Z, et al. Performance analysis of neural network architectures for time series forecasting: A comparative study of RNN, LSTM, GRU, and hybrid models [J]. MethodsX, 2025, 15: 103462.
[7] T, S., & G, M. R. (2025). Accurate hourly AQI prediction using Temporal CNN-LSTM-MHA+GRU: A case study of seasonal variations and pollution extremes in Visakhapatnam, India. Results in Engineering, 106303. https:

DOI: 10.54254/2754-1169/2025.BL27522

//doi.org/10.1016/j.rineng.2025.106303
[8] Xie J, Liu F, Liu S, et al. An Approach to Spatiotemporal Air Quality Prediction Integrating SwinLSTM and Kriging Methods [J]. Sustainability, 2025, 17(7): 2918.
[9] Kul, E. (2021). Shanghai Air Pollution and Weather 2014–2021 [Dataset]. Kaggle. https: //www.kaggle.com/datasets/erhankul/shanghai-air-pollution-and-wheather-20142021