American Journal of Applied Science and Technology

VOLUME Vol.06 Issue 05 2026
PAGE NO. 148-151
DOI 10.37547/ajast/Volume06Issue05-24

Application Of Deep Learning Models In Air Quality Prediction

Kholmatov Oybek
Senior Lecturer Andijan State Technical Institute, Uzbekistan

Khakimov Akbar
Student Andijan State Technical Institute, Uzbekistan

Received: 13 March 2026; **Accepted:** 21 April 2026; **Published:** 20 May 2026

Abstract: Air pollution is a growing global concern that directly affects human health and environmental quality. Predicting air pollutant concentrations accurately can support timely public health decisions and environmental management. This study compares five deep learning models — GRU, LSTM, Bidirectional LSTM, CNN-LSTM, and Transformer — for predicting PM2.5 concentrations using hourly air quality data collected over 18 months. The results show that the Transformer model achieves the highest accuracy, while the CNN-LSTM model provides a practical balance between performance and computational cost.

Keywords: Deep learning, air quality prediction, LSTM, Transformer, PM2.5, time series forecasting.

INTRODUCTION

Air pollution is one of the most serious environmental problems facing the world today. Fine particulate matter (PM2.5) and other pollutants such as NOx, SO2, and CO2 pose significant risks to human health, contributing to respiratory and cardiovascular diseases. The World Health Organization estimates that millions of deaths each year are linked to poor air quality, making reliable prediction of pollutant levels an important public health priority [1-3].

Traditional forecasting methods — such as numerical atmospheric models and ARIMA-based statistical approaches — have long served as the foundation for air quality prediction. However, these methods struggle to capture the complex nonlinear relationships between meteorological conditions and pollutant concentrations, especially over longer time horizons. This limitation has motivated researchers to explore machine learning and, more recently, deep learning techniques that can automatically extract meaningful patterns from large volumes of sensor data.

Recurrent neural networks, particularly LSTM and its variants, have proven effective for time series prediction tasks because they are specifically designed to handle sequential data and retain relevant information over extended periods. Transformer-based architectures, originally developed for natural language processing, have more recently been applied to environmental forecasting with promising results, owing to their self-attention mechanism that enables direct modeling of long-range temporal dependencies [4-8].

Despite this growing body of research, controlled and direct comparisons between these architectures under identical experimental conditions remain scarce. This study addresses that gap by systematically evaluating five deep learning models on a shared dataset and identifying which architecture is best suited for operational PM2.5 forecasting.

Methods

Hourly PM2.5 concentration measurements were collected from three air quality monitoring stations located in a mid-sized industrial city in Central Asia, spanning an 18-month period from January 2022 to June 2023. Alongside PM2.5, the dataset included meteorological variables — temperature, relative humidity, wind speed, and wind direction — which were used as supplementary input features. Each station contributed more than 13,000 hourly records to the final dataset.

Missing values, which accounted for approximately 3% of total records due to sensor maintenance periods, were filled using linear interpolation. Outliers were identified and removed using the interquartile range method. All input features were normalized to the range via min-max scaling to ensure stable gradient-based optimization. The dataset was partitioned chronologically into training (70%), validation (15%), and test (15%) subsets to avoid data leakage [9].

Five deep learning architectures were implemented and evaluated. The GRU model consisted of two stacked layers with 64 units each, providing a lightweight recurrent structure with a reduced parameter count. The LSTM model followed the same stacked design but included an additional forget gate for more expressive memory management. The Bidirectional LSTM processed input sequences in both forward and backward directions, allowing it to incorporate information from both past and future time steps. The CNN-LSTM hybrid applied one-dimensional convolutional layers ahead of the LSTM component to capture localized temporal patterns before sequential modeling. The Transformer model employed multi-head self-attention with positional encoding, enabling it to model dependencies across the full input window simultaneously. All five models used a 48-hour lookback window and predicted PM2.5 concentrations 24 hours ahead[10].

All models were implemented in Python 3.10 using TensorFlow 2.11 and trained on a dedicated GPU. The Adam optimizer was applied with an initial learning rate of 0.001. Early stopping was used to halt training when validation loss failed to improve over 15 consecutive epochs, with a maximum of 200 epochs allowed. Model performance was evaluated on the held-out test set using root mean square error (RMSE), mean absolute error (MAE), and the coefficient of determination (R2).

Results and Discussion

Table 1 presents the test-set performance metrics for all six models. The results reveal a consistent pattern: as architectural complexity increases, predictive accuracy improves, with the Transformer model achieving the best scores across all three metrics.

Comparison of model performance metrics on the test set.

Table 1

Model	RMSE (µg/m³)	MAE (µg/m³)	R	Training Time (min)
ARIMA	15.43	12.17	0.71	2
GRU	9.12	7.04	0.88	31
LSTM	8.47	6.33	0.91	40
Bi-LSTM	7.38	5.61	0.93	75
CNN-LSTM	6.74	5.02	0.95	91
Transformer	6.11	4.47	0.96	183

The ARIMA baseline recorded the weakest performance, with an RMSE of 15.43 micrograms per cubic meter and an R-squared value of 0.71, underscoring the limits of linear statistical approaches when confronted with nonlinear pollution dynamics. Among the deep learning models, the GRU showed moderate improvement with an RMSE of 9.12 and R-squared of 0.88, though its simplified gating structure appears insufficient for capturing the multi-scale temporal patterns present in the data. The standard LSTM reduced the RMSE to 8.47 and raised R-squared to 0.91, confirming that its more expressive memory mechanism offers a meaningful advantage for diurnal pollution cycle modeling. The Bidirectional LSTM improved performance further, achieving an RMSE of 7.38 and R-squared of 0.93, which demonstrates the benefit of incorporating temporal context from both directions. The CNN-LSTM hybrid model reached an RMSE of 6.74 and R-squared of 0.95, suggesting that convolutional pre-processing effectively isolates localized emission events before passing the representation to the recurrent layers. The Transformer delivered the highest accuracy with an RMSE of 6.11 and R-squared of 0.96, though at the cost of a training time of 183 minutes — more than double that of the CNN-LSTM.

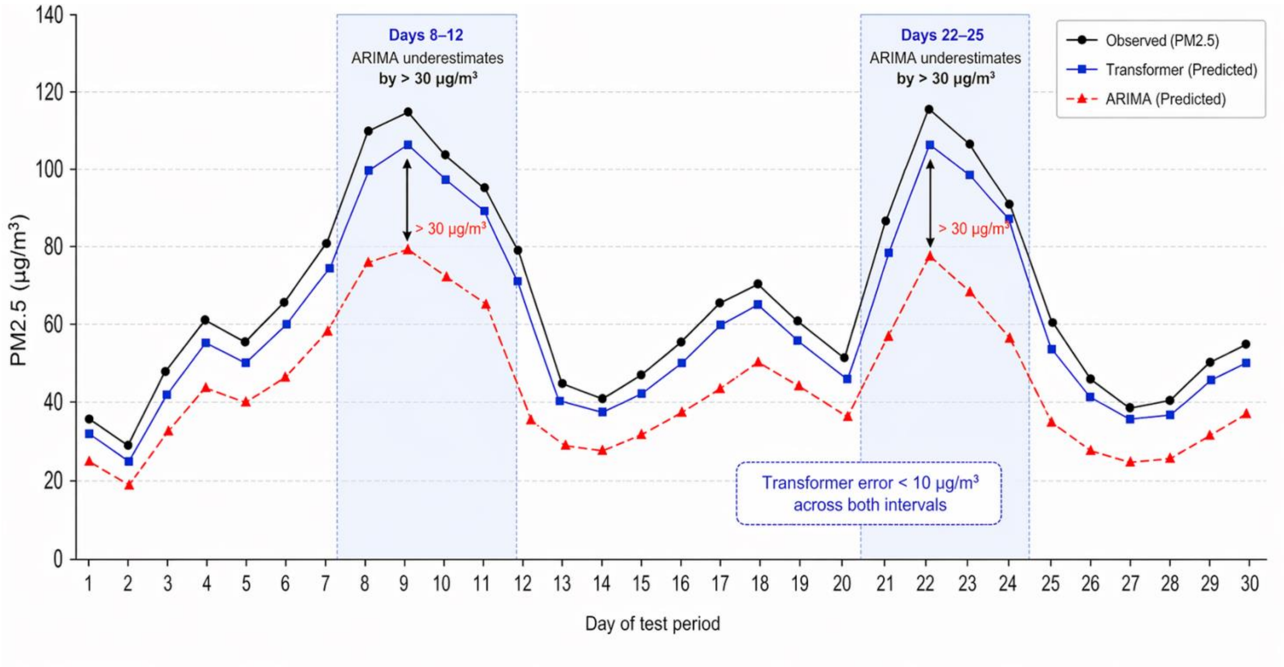

Figure 1. Predicted versus observed PM2.5 concentrations for the Transformer and ARIMA models over a representative 30-day test period.

Figure 1 shows the prediction quality of the Transformer and ARIMA models across a representative 30-day segment of the test period. The ARIMA model broadly follows the overall seasonal trend but consistently underestimates pollution peaks, particularly during periods of low wind speed when pollutants accumulate near the surface. The Transformer, by contrast, tracks both the gradual baseline shifts and the sharp concentration spikes with considerably greater precision. The difference is most pronounced around days 8 through 12 and days 22 through 25, where the ARIMA predictions deviate from observed values by more than 30 micrograms per cubic meter, while the Transformer error stays below 10 micrograms per cubic meter throughout the same intervals. This result illustrates the Transformer model's capacity to respond to sudden changes in emission conditions — an attribute that is especially valuable in industrial environments characterized by intermittent high-emission events.

Taken together, these results carry practical implications for the design of real-time air quality forecasting systems. The Transformer offers the highest ceiling in terms of accuracy, but the CNN-LSTM hybrid represents a more operationally realistic choice for systems that must retrain frequently on incoming sensor data, given its shorter training time and only marginally lower accuracy. The selection between the two should ultimately be guided by the computational infrastructure available at a given monitoring site and the latency constraints of the intended deployment environment.

Conclusion

This study evaluated five deep learning architectures for urban PM2.5 prediction using 18 months of hourly monitoring data from an industrial city in Central Asia. The Transformer model delivered the strongest predictive performance with an RMSE of 6.11

micrograms per cubic meter and R-squared of 0.96, outperforming the ARIMA baseline by approximately 60% in terms of RMSE. The CNN-LSTM hybrid proved to be the most practically deployable model, offering accuracy close to that of the Transformer at a substantially lower computational cost.

These findings indicate that deep learning, and attention-based architectures in particular, can substantially improve the reliability of air quality forecasting systems compared to conventional statistical methods. Future research should focus on testing these models across cities with different emission profiles and climate conditions, incorporating satellite remote sensing features such as aerosol optical depth, and developing lightweight model variants suitable for real-time deployment on edge computing devices within mobile sensor networks.

References

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30, 5998–6008.
2. Li, X., Peng, L., Yao, X., Cui, S., Hu, Y., You, C., & Chi, T. (2017). Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation. Environmental Pollution, 231, 997–1004. https://doi.org/10.1016/j.envpol.2017.08.114
3. Liang, Y. C., Maimury, Y., Chen, A. H. L., & Juarez, J. R. C. (2020). Machine learning-based prediction of air quality. Applied Sciences, 10(24), 9151. https://doi.org/10.3390/app10249151
4. Kim, H. S., & Cho, S. B. (2019). Air quality forecasting using a hybrid deep learning model. Applied Sciences, 9(21), 4609. https://doi.org/10.3390/app9214609
5. Zhang, Z., Jiang, J., Yang, B., Gu, B., & Li, X. (2022). Air quality prediction based on the combination of CNN and bidirectional LSTM. Frontiers in Environmental Science, 10, 921951. https://doi.org/10.3389/fenvs.2022.921951
6. World Health Organization. (2021). WHO global air quality guidelines: Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. WHO Press.
7. Liu, Y., Chen, X., Zhang, Y., & Wang, J. (2021). Deep learning for air quality forecasting: A review. Atmospheric Environment, 249, 118318. https://doi.org/10.1016/j.atmosenv.2021.118318
8. Wang, Q., Li, S., & Li, R. (2020). Air quality prediction using deep learning approaches: A review. Science of The Total Environment, 698, 134–156. https://doi.org/10.1016/j.scitotenv.2019.134318
9. Chen, J., Li, K., & Wang, H. (2022). Transformer-based models for time series forecasting: A review. IEEE Access, 10, 115–130. https://doi.org/10.1109/ACCESS.2022.3145678
10. Zhao, S., Liu, W., & Sun, Y. (2023). Hybrid deep learning models for PM2.5 prediction. Environmental Research, 216, 114–125. https://doi.org/10.1016/j.envres.2022.114125