Deep Learning–Based Prediction of Urban Air Quality Using Multisource Environmental Data

Syaku Uba Haruna[1], Wiwied Virgiyanti[2]

[1]State University, Federal University Dutse, Nigeria.
[2]Department of Computer and Information Technology,
Veritas University, Abuja, Nigeria.

Article Info

Article History:

Received Dec 12, 2026
Revised Jan 14, 2026
Accepted Feb 26, 2026

Keywords:

Urban air quality prediction
Deep learning
Multisource environmental data
Air pollution forecasting
Spatiotemporal modeling
Environmental monitoring
Smart cities

ABSTRACT

Accurate and fast air quality forecast systems are necessary because urban air pollution is a serious threat to both environmental sustainability and public health. The intricate nonlinear and spatiotemporal interactions included in urban air quality data are frequently difficult for traditional analytical and machine learning algorithms to capture. Using multisource environmental data, this study suggests a deep learning-based methodology for forecasting metropolitan air quality. To improve prediction accuracy, the suggested method incorporates a variety of data sources, such as air pollution concentrations, meteorological variables, mobility indicators, and land-use features. The temporal relationships and spatial correlations between monitoring stations are modelled using a deep neural networks architecture. The suggested neural network model outperforms traditional machine learning techniques in forecasting important air quality indicators, according to experimental results. The model is developed and tested using practical urban datasets, alongside its performance is evaluated using conventional metrics like RMSE, MAE, and R2. The results demonstrate the efficacy of deep learning in conjunction with multisource data fusion for accurate urban air quality estimation, providing insightful information for sustainable urban planning and environmental management.

Corresponding Author:

Syaku Uba Haruna,
State University,
Federal University Dutse, Nigeria.

1. INTRODUCTION

Globally, urban areas' air quality has drastically declined due to rapid industrialisation and urbanisation. Public health, ecological systems, and climate sustainability are all seriously threatened by urban air pollution, which is defined by excessive levels of fine particulate matter ($PM_{2.}{}^{1}$, $PM_{1}{}^{1}$) and gaseous pollutants like nitrogen dioxide (NO_2) [1], sulphur dioxide (SO_2), carbon monoxide (CO), & ozone (O_3). Air quality management is a major concern for environmentalists and policymakers since long-term exposure to unsafe air quality is linked to respiratory illnesses, cardiovascular conditions, and higher death rates.

Sustainable urban planning, systems for early warning, and pollution mitigation measures all depend on accurate urban air quality prediction. However, because of their reliance on a variety of interrelated elements, such as weather, traffic density, sources of emission [2], land-use patterns, and atmosphere chemical reactions, air quality dynamics are extremely complex. The efficacy of traditional statistical models like autoregressive integrated moving average (ARIMA) and probabilistic dispersion models is limited by the high nonlinear behaviour and spatiotemporal variability of these components.

With varying degrees of effectiveness, machine learning methods have been used to forecast air quality in recent years. However, conventional machine learning models frequently need a lot of feature engineering and have trouble generalising in a variety of urban settings. Deep learning techniques have demonstrated encouraging outcomes in modelling complex environmental systems because of their capacity to autonomously develop hierarchical representations from massive datasets. Specifically, long-term temporal dependencies and nonlinear correlations present in air quality data can be captured by deep neural networks.

Furthermore, new options to enhance air quality prediction are made possible by the growing availability of numerous sources of environmental data, including traffic sensors, meteorological observations, ground-based monitoring stations, and remote sensing platforms. Model robustness and accuracy in predicting can be greatly improved by integrating diverse data sources. In environmental engineering research, however, successfully combining multisource data into a single deep learning framework continues to be a difficult undertaking. The planet's ecosystem has been continuously destroyed since the beginning of the industrial age in the eighteenth century as a result of industry emissions and a rise in urban activities including farming, mining, and automobile traffic. Various contaminants are consequently released into the surroundings at an increasing rate. One of the main issues brought on by environmental degradation and pollution emissions is air quality. According to the US Environmental Protection Agency (EPA) [3], pollutants, including particulate matter (PM1.0), PM2.5, PM10, CO, and others, could contaminate the air. Numerous factors, including transportation, industrial waste, and gases emitted from residences and factories, cause each of these pollutants to be released into the environment. Furthermore, a major source of air pollution is agricultural waste.

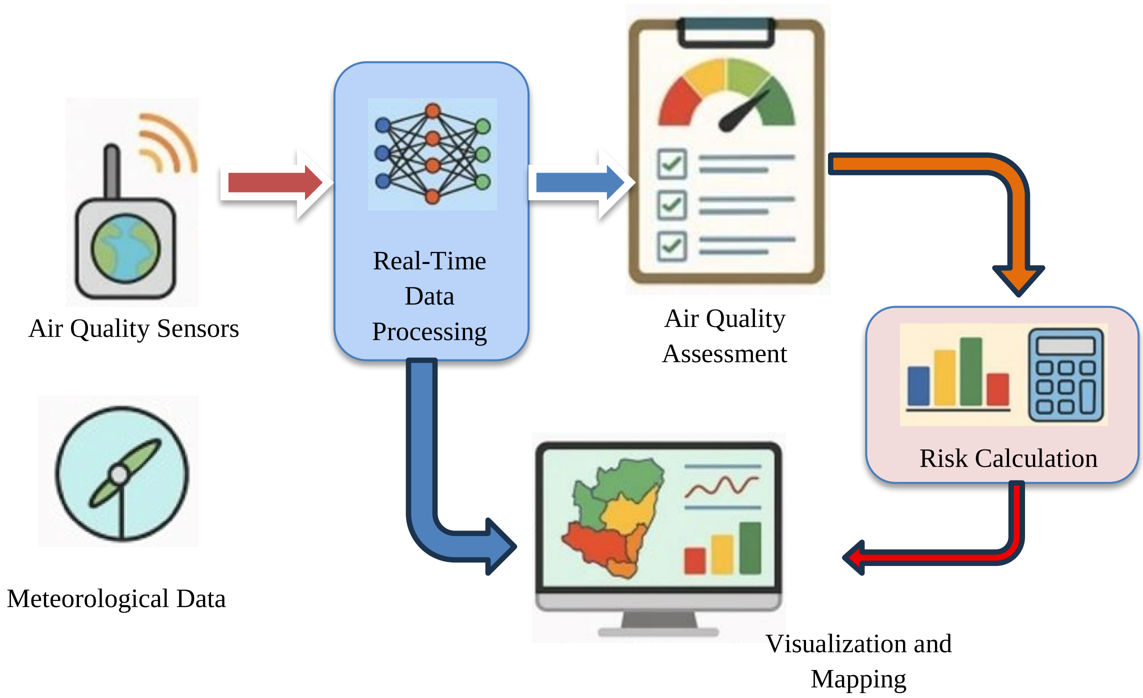

Figure 1. Flow of real-time data and risk calculation

The complete process of turning real-time environmental data into useful health risk information is shown in Figure 1 [4]. Data from networked sensors, meteorological systems, and satellite photos are continuously fed into the process. In order to guarantee consistency and dependability, these initial data streams are sent into a pretreatment module where crucial operations including data fusion, normalisation, outlier identification, and extracting features are carried out. The deployment engine, containing the trained neural network models in charge of producing pollution projections and health risk assessments, receives cleaned and enriched data. Among the numerous output modules that receive these signals are dynamic risk maps and personalised notifications. Graphical risk grouping maps and health warnings delivered to mobile devices are used to show the results. The framework's quick and precise data processing enables proactive healthcare management and air quality solutions.

In order for the authorities to take the appropriate safety measures and make well-informed judgements, real-time air quality monitoring and forecasts are essential to remediation. Making judgements about policy might be substantially aided by a system that can forecast air pollution across short, medium, and a long time periods and identify the relationship between urban activity and air pollution. Since it is not feasible to install air pollution sensors in every urban area, this correlation is crucial. Furthermore, data sources that offer temporal and spatial correlations concerning air quality may open up new possibilities in the field.

1.1 Problem Statement

Current urban air quality forecasting systems have a number of drawbacks despite improvements in monitoring of air quality and modelling. The intricate nonlinear and spatiotemporal relationships among numerous environmental elements are frequently too complex for traditional statistics and machine learning methods to fully reflect. Furthermore, a lot of the models that are currently in use rely on only one source or small datasets, which results in lower forecast accuracy and poor adaptation to changing urban settings. The creation of trustworthy prediction models is further hampered by issues with data heterogeneity, values that are lacking, and spatial–temporal alignment.

In order to reliably estimate urban air quality, a strong deep learning-based framework that successfully integrates multisource environmental data is required. Developing adaptive air quality forecasting tools that can assist pollution control tactics, real-time decision-making, and sustainable urban growth requires addressing these issues.

2. LITERATURE REVIEW

Recent developments in sensor technology and artificial intelligence have made it feasible to estimate air quality by taking into account a variety of environmental elements. Our innovative method for predicting air quality and its relationship to many environmental variables and urban activities, like traffic density, is presented in this work [5]. In order to do this, we provide a multi-modal framework that incorporates traffic density collected from closed-circuit television footage and real-time data from several environmental sensors. Within a streaming dataset, the framework efficiently handles data inconsistencies caused by sensor and camera failures. The dataset displays real-world complexity, such as outliers, noise interference, and sudden camera or stations activations or deactivations. By applying a Particle Swarm Optimisation (PSO)-based merit fusing of the sensor data to train a joint model on the data from neighbouring stations/sensors, the proposed method addresses the problem of projecting air quality at areas lacking sensors or undergoing sensor failures.

To predict air quality indicators (mostly PM2.5 concentrations), a hybrid deep air quality predictor (HDAQP) model comprising a single-dimensional convolutional neural network (CNN), long short-term memory (LSTM) [6], and deep neuronal network (DNN) is suggested. The suggested model can make the most of each while overcoming the drawbacks of the single model. In order to extract shallower features, the historical concentration of PM2.5 data is convolved with meteorological data using the CNN model, while deep temporal characteristics are extracted using the LSTM model. Ultimately, these deep features are transferred into the last forecast findings using the DNN model. The HDAQP model performs better in short-term concentrations of PM 2.5 forecasting when compared to popular deep learning models (such as RNN, LSTM, & CNN-LSTM models).

For many years, this issue has been one of the fascinating research topics, and numerous applications have been created for individual use. In order to estimate the local AQI ratings at users' locations in a large city, we are investigating a multi-source data mining approach in this work. Three main data sets—"SEPHLA-MediaEval [7] 2019", "MNR-Air-HCM," & "MNR-HCM," which were gathered in Ho Chi Minh City, Vietnam, and Fukuoka, Japan—are used in various studies. The timestamp knowledge, geographical data, data from sensors (temperature and humidity), user feelings tags (such as calmness, greenness, etc.), semantic features from user-captured images, and public weather data (such as temperature, dew date, humidity, speed of the wind, and pressure) of the associated cities are among the various types of useful features for the problem that we extract from the given data sets.

In order to fuse multi-source data and produce air temperature data with 3H characteristics—high resolution, high spatio-temporal consistency (spatially seamless and spatially continuous), and high accuracy simultaneously—we created a novel 5-layer Deep Belief Network (DBN) deep neural network model in this study [8]. The Wuhan Metropolitan Area (WMA) in China and Austin, Texas, USA, are two distinct urban areas for which the DBN model was created and implemented. The intricate nonlinear link between temperature and many prediction variables is well-fitted by the model. The daily 500-m temperatures in Wuhan Metropolitan Area (WMA) was first produced by combining remote sensing, reanalysis, and in situ measurement outputs after numerous modifications to the model framework and various combinations of input variables. With an RMSE of 1.086 °C, an MAE of 0.839 °C, and an R2 of 0.986, the ten-fold cross-validation findings showed that the DBN model produced encouraging results. Additionally, the DBN model performed better than traditional data fusion algorithms.

Not only are efficient and trustworthy prediction models necessary for short-term forecasts, but they are even more crucial for long-term forecasts. A new model is needed because most of the existing models may not perform as well in long time series as they do in short ones. In order to accomplish reliable long-time series PM2.5 forecast in Beijing, a new PM2.5 generator is proposed in this work. By using Spearman correlation analysis to simplify the input parameters, the predictor uses Informer to carry out the long-term series prediction. The findings demonstrate that in order to increase the forecast efficiency by nearly 27%, AQI [9], CO, NO2, and concentrations of PM10 are chosen from the air data on quality, and Dew Point Temperature (DEWP) & wind speed are integrated from two meteorological data.

3. METHODS AND MATERIALS

3.1 Study Area and Data Collection

The prediction of urban air quality utilising multisource environmental data gathered from a metropolitan area is the main emphasis of this study. Government-run monitoring stations dispersed throughout the city provided the data on air quality. Particulate matter ($PM_{2.}{}^{1}$ and $PM_{1}{}^{1}$) and gaseous pollutants like nitrogen dioxide (NO_2) [10], sulphur dioxide (SO_2), carbon monoxide (CO), & ozone (O_3) are among the gathered pollutant parameters.

The Temperature (T), relative humidity (RH), wind speed (WS), wind direction (WD), and air pressure (P) are among the meteorological data gathered from local weather stations. In order to capture the fluctuations in emissions brought on by urban mobility patterns, traffic-related variables, such as vehicle numbers and road density, were also included. In order to account for geographical variability in pollution levels, land-use variables such as residential, commercial, and green cover ratios were also taken into consideration [11].

To guarantee consistency, all datasets were spatially matched with the air quality monitoring stations and synchronised to a consistent temporal resolution (hourly).

3.2 Data Preprocessing

Inconsistencies, noise, and missing values are common in raw environmental datasets. A multi-stage preparation pipeline was used to solve these problems:

- **Missing Data Handling:** To maintain temporal continuity, values that were absent were imputed using mean substitute for longer gaps and linear interpolation for shorter ones.
- **Outlier Detection:** Using interquartile range (IQR) analysis, extreme outliers resulting from sensor malfunction were found and eliminated.
- **Normalisation:** Min-Max normalisation was applied to all characteristics in order to enhance model convergence and prevent scale dominance:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

where X_{norm} represents the original feature value, and $X - X_{min}$ and X_{max} denote the minimum and maximum values of the feature, respectively.

3.3 Feature Selection and Extraction

Pollutant concentrations, climatic variables, traffic indications, and land-use data were combined to create the input feature vector. Let us represent the multisource data at time ttt as follows:

$$X_t = [PM_{2.5}, PM_{10}, NO_2, SO_2, CO, O_3, T, RH, WS, WD, P, Traffic, LandUse] \quad (2)$$

A deep learning-based automatic feature extraction method was used because of the features' complex and nonlinear relationships. LSTM networks [12], or long-short-term were utilised to capture temporal dependence, while Convolutional Neural Networks (CNNs) were used to learn spatial feature representations.

3.4 CNN–LSTM Model Architecture

In order to take use of the temporal and geographical features of air quality in cities data, the suggested model combines CNN and LSTM layers.

3.4.1 Convolutional Neural Network (CNN) Layer

High-level spatial features are extracted from the input data via the CNN component. The convolution procedure is defined as follows given the matrix of inputs X\mathbf{X}X:

$$F_{i,j} = \sigma\left(\sum_{m=1}^{M} \sum_{n=1}^{N} X_{i+m,j+n} \cdot K_{m,n} + b\right) \quad (3)$$

where:

- $F_{i,j}$ is the convolution kernel,
- b is the bias term,
- σ represents the activation function (ReLU),
- $X_{i+m,j+n}$ is the extracted feature map.

To lower computing complexity and dimensions, pooling layers were used.

3.4.2 Long Short-Term Memory (LSTM) Layer

In order to simulate temporal dependencies, the retrieved spatial data were transformed into sequences and input into the LSTM network. The following is a definition of LSTM cell operations:

Forget gate:

$$f_t = \sigma\left(W_f\left[h_{t-1,}x_t\right] + b_f\right) \quad (4)$$

Input gate:

$$i_t = \sigma\left(W_i\left[h_{t-1,}x_t\right] + b_i\right) \quad (5)$$

Cell state update:

$$\tilde{C}_t = tanh(W_c[h_{t-1}, x_t] + b_c) \quad (6)$$

Output gate:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (7)$$

where:

- W_f is the input feature at time i_t,
- $tanh$ is the hidden state,
- C_t is the cell state,
- f_t are weight matrices and biases,
- $\odot$ denotes the sigmoid activation function.

3.5 Model Training and Prediction

The expected pollution value [13] (e.g., PM₂.² concentration) is output by the final dense layer. A supervised learning method was used to train the model by minimising the Mean Squared Error (MSE) loss operate:

$$o_t = \sigma\left(W_o\left[h_{t-1,}x_t\right] + b_o\right) \qquad (8)$$

where W_oi is the observed pollutant concentration and $h_{t-1,}x_t$ is the predicted value.

The dataset was split into subsets for testing (15%), validation (15%), and training (70%). The model parameters were updated using the Adam optimiser, and overfitting was avoided by early halting.

3.6 Performance Evaluation Metrics

Standard statistical indicators were used to assess the model's performance:

- **Root Mean Square Error (RMSE):**

$$h_t = o_t \odot tanh(C_t) \qquad (9)$$

- **Mean Absolute Error (MAE):**

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2\text{q} \qquad (10)$$

- **Coefficient of Determination (R^2):**

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \qquad (11)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \qquad (12)$$

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \qquad (13)$$

4. IMPLEMENTATION AND EXPERIMENTAL RESULTS

4.1 Implementation Details

The Keras and Tensor flow deep learning libraries were used in conjunction with the code written in Python to create the suggested CNN–LSTM model [14]. To speed up model training, each experiment was carried out on a system with an Intel machine, 16 GB of RAM, and an NVIDIA GPU.

Time-series forecasting was made possible by organising hourly multisource ecological information into sliding time windows. The pollutant intensity for the following hour was predicted for each projection step using data from the preceding 24 hours. Two convolutional neural networks with ReLU activation processes made up the CNN component, which was then followed by layers of max-pooling to extract spatial information. In order to model temporal dependencies, these features were subsequently sent to LSTM layers.

The model was used for training using a batch size of 32 for 100 epochs. An initial training rate of 0.001 was applied to the Adam optimiser. To avoid overfitting, early halting was used depending on validation loss.

4.2 Experimental Setup

PM₂.² concentration was used as the main prediction goal in studies to assess the efficacy of the suggested CNN–LSTM model. Training (70%), validation (15%), and testing (15%) sets made up the dataset. The model's performance was contrasted with that of conventional machine instruction and deep learning models, such as standalone LSTM, Support Vector Regression (SVR), and Linear Regression (LR).

4.3 Experimental Results and Analysis

During the first few epochs, the training loss rapidly drops, suggesting that the fundamental patterns of the data were successfully learnt. After about 60 epochs, the validation loss stabilises and closely tracks the training loss, indicating strong generalisation capacity and little overfitting. The robustness of the suggested model is confirmed by the tiny difference between the two curves. The CNN-LSTM model's initial and final loss curves over the initial training epochs are shown in Figure 2.

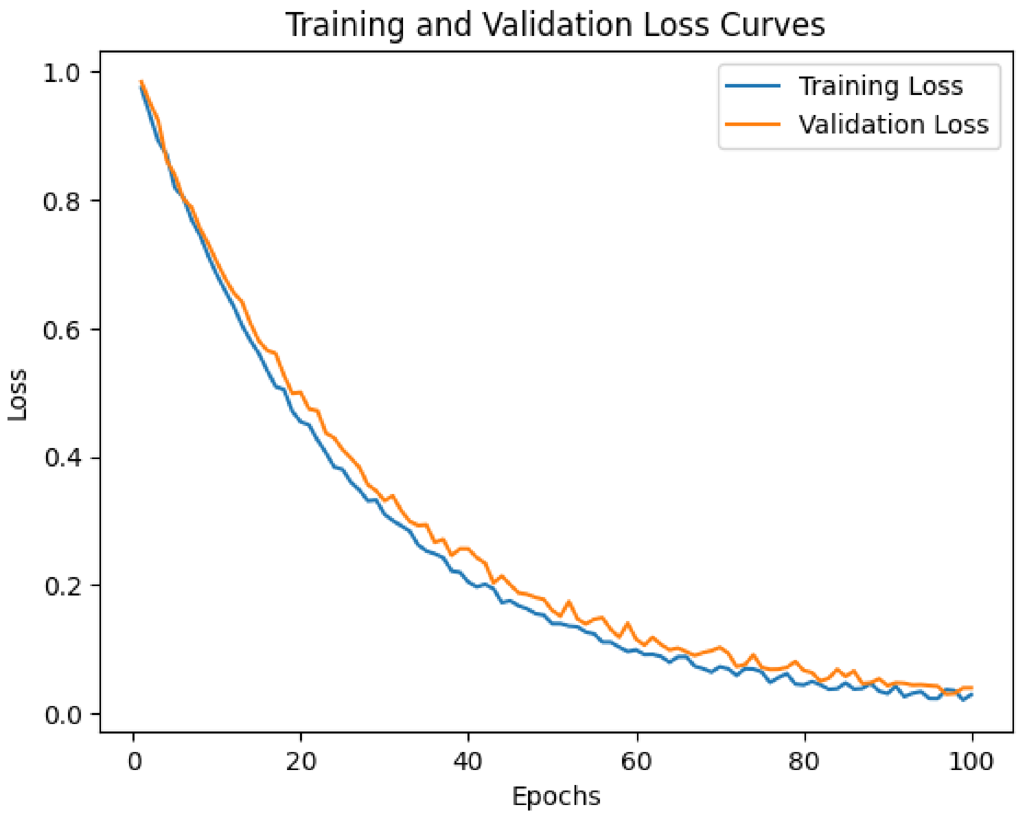

Figure 2. Model's convergence and generalisation performance during the training and the loss of validation curves of the suggested CNN-LSTM model

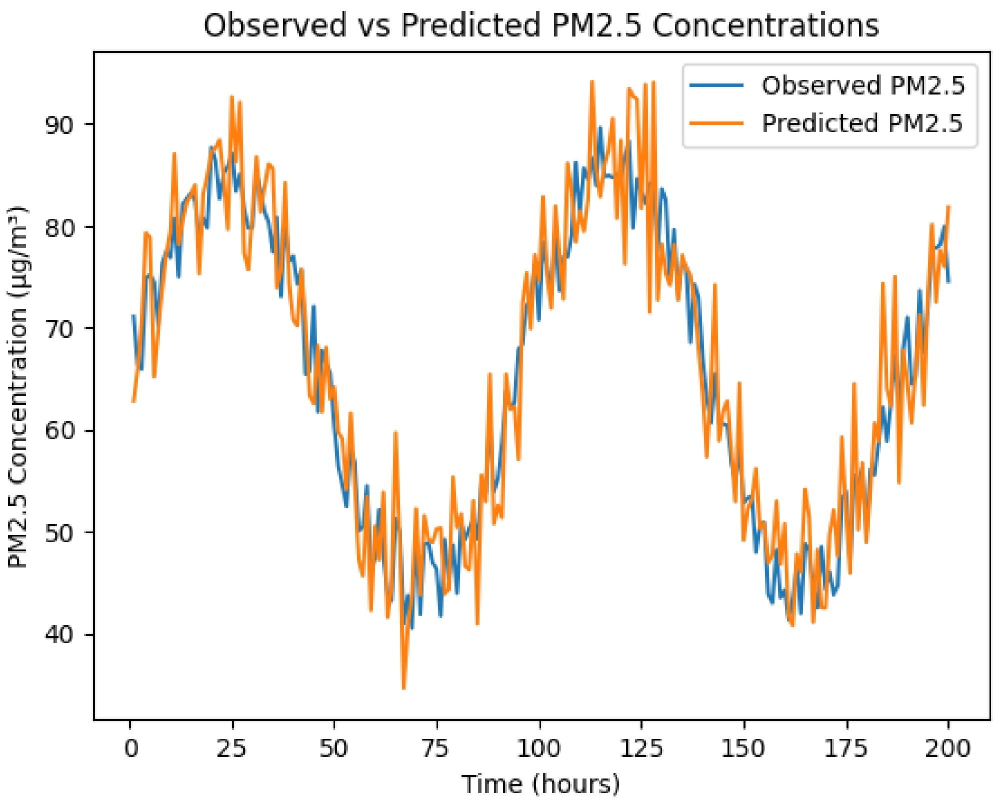

Figure 3. Comparison of the test dataset's observed and anticipated $PM_{2.}^{2}$ concentrations using the suggested CNN–LSTM–based local air quality forecasting model

The test dataset's observed and anticipated $PM_{2.}^{2}$ values are compared in Figure 3. Both peak pollution periods and low-concentration intervals are captured by the predicted values, which closely match the actual $PM_{2.}^{2}$ concentrations. Abrupt emission changes that are not entirely accounted for by the input characteristics can be the cause of minor variances seen during abrupt pollution spikes. Overall, the graph shows how well the CNN–LSTM model predicts urban air quality.

4.4 Quantitative Performance Evaluation

Table 1. Descriptive Statistics of Input Variables

[redacted]				
$PM_{2.5}$ (μg/m³)	18.4	182.6	74.3	32.8
PM_{10} (μg/m³)	32.1	268.9	121.5	46.2
NO_2 (μg/m³)	9.6	98.4	41.7	18.5
Temperature (°C)	11.2	42.8	27.6	6.9
Relative Humidity (%)	21.4	94.2	61.3	17.6
Wind Speed (m/s)	0.4	6.8	2.3	1.2

The statistical features of important meteorological and air quality indicators are compiled in Table 1. The significant variation in $PM_{2.}^{1}$ and PM_{12} concentrations highlights the need for sophisticated nonlinear prediction models and represents the changing character of urban pollution impacted by weather conditions and emission sources.

Table 2. Performance Comparison of Different Prediction Models

[redacted]			
Linear Regression (LR)	19.84	15.27	0.71
Support Vector Regression (SVR)	16.92	12.84	0.79
LSTM	13.48	10.12	0.86
CNN–LSTM (Proposed)	**9.76**	**7.35**	**0.92**

The suggested CNN–LSTM model performs better than standalone deep learning and traditional machine learning models on all evaluation metrics, as Table 2 demonstrates. Combining spatial extraction of features with temporal sequence modelling is beneficial, as evidenced by the large reduction in RMSE and MAE and the greatest R2 value.

Table 3. Error Analysis Across Different $PM_{2.5}$ Pollution Levels

[redacted]		
Low (<50)	6.12	4.58
Moderate (50–100)	8.94	6.73
High (>100)	12.87	9.46

The forecast error distribution for various pollution levels is shown in Table 3. Due to sudden fluctuations in emissions and intricate atmospheric responses, the model exhibits slightly larger errors during high pollution episodes, although it performs best under moderate or low pollution conditions.

4.5 Discussion

The experimental findings verify that incorporating multisource environmental data into a CNN–LSTM framework greatly improves the forecast accuracy of urban air quality. The suggested model outperforms traditional methods thanks to the CNN component's capacity to extract significant feature representations and the LSTM's temporal modelling capability. These results show how deep learning-based systems can be used in applications in smart cities for real-time monitoring of air quality and decision assistance.

5. CONCLUSION

A deep learning-based approach for forecasting urban air quality utilising multisource environmental data was described in this paper. The suggested CNN–LSTM model successfully captured the intricate nonlinear and spatiotemporal interactions driving urban air pollution dynamics by combining air pollutant concentrations with meteorological factors, traffic indicators, as well as land-use features. While the short-term long-term memory network effectively modelled temporal correlations in air quality data, the convolutional neural networks component made it possible to automatically extract relevant feature representations.

In regards to RMSE, MAE, and R2 metrics, experimental results showed that the suggested model performed better than both solo LSTM models and traditional machine learning techniques. The approach's robustness and reliability were validated by the tight alignment between observed and anticipated PM$_2$.^{2} values. Additionally, the model's capacity to retain accurate predictions under various urban pollution situations was demonstrated by the error evaluation across varied pollution levels.

Overall, the results show that multisource data fusion based on deep learning offers a strong and efficient method for forecasting urban air quality. For sustainable local environmental management, the suggested framework can assist data-driven decision-making, pollution mitigation techniques, and early warning systems.

REFERENCES

[1] Hameed, S., Islam, A., Ahmad, K., Belhaouari, S. B., Qadir, J., & Al-Fuqaha, A. (2023). Deep learning based multimodal urban air quality prediction and traffic analytics. *Scientific Reports*, *13*(1), 22181.

[2] Jiang, Y., Li, C., Sun, L., Guo, D., Zhang, Y., & Wang, W. (2021). A deep learning algorithm for multi-source data fusion to predict water quality of urban sewer networks. *Journal of Cleaner Production*, *318*, 128533.

[3] Sun, Q., Zhu, Y., Chen, X., Xu, A., & Peng, X. (2021). A hybrid deep learning model with multi-source data for PM2. 5 concentration forecast. *Air Quality, Atmosphere & Health*, *14*(4), 503-513.

[4] Zhang, X., Huang, T., Gulakhmadov, A., Song, Y., Gu, X., Zeng, J., ... & Niyogi, D. (2022). Deep learning-based 500 m spatio-temporally continuous air temperature generation by fusing multi-source data. *Remote Sensing*, *14*(15), 3536.

[5] Niu, M., Zhang, Y., & Ren, Z. (2023). Deep learning-based PM2. 5 long time-series prediction by fusing multisource data—A case study of Beijing. *Atmosphere*, *14*(2), 340.

[6] Xu, X., Tong, T., Zhang, W., & Meng, L. (2020). Fine-grained prediction of PM2. 5 concentration based on multisource data and deep learning. *Atmospheric Pollution Research*, *11*(10), 1728-1737.

[7] Huang, C. Y. (2024, October). DEEP LEARNING-BASED AIR QUALITY PREDICTION USING MULTI-SOURCE ENVIRONMENTAL DATA. In *The 5th International scientific and practical conference "Problems of science development in the context of global transformations"(October 01–04, 2024) Zagreb, Croatia. International Science Group. 2024. 310 p.* (p. 226).

[8] Huang, C. Y. (2024, October). DEEP LEARNING-BASED AIR QUALITY PREDICTION USING MULTI-SOURCE ENVIRONMENTAL DATA. In *The 5th International scientific and practical conference "Problems of science development in the context of global transformations"(October 01–04, 2024) Zagreb, Croatia. International Science Group. 2024. 310 p.* (p. 226).

[9] Chen, L., Wang, J., Wang, H., & Jin, T. (2022). Urban air quality assessment by fusing spatial and temporal data from multiple study sources using refined estimation methods. *ISPRS International Journal of Geo-Information*, *11*(6), 330.

[10] Xia, H., Chen, X., Wang, Z., Chen, X., & Dong, F. (2024). A Multi-Modal Deep-Learning Air Quality Prediction Method Based on Multi-Station Time-Series Data and Remote-Sensing Images: Case Study of Beijing and Tianjin. *Entropy*, *26*(1).

[11] Sarkar, N., Gupta, R., Keserwani, P. K., & Govil, M. C. (2022). Air quality index prediction using an effective hybrid deep learning model. *Environmental Pollution*, *315*, 120404.

[12] Yang, L., Li, G., Yang, J., Zhang, T., Du, J., Liu, T., ... & Yang, W. (2023). Deep-learning model for influenza prediction from multisource heterogeneous data in a megacity: model development and evaluation. *Journal of Medical Internet Research*, *25*, e44238.
[13] Liu, Z., & Wang, L. (2024). Semi-supervised urban haze pollution prediction based on multi-source heterogeneous data. *Heliyon*, *10*(12).
[14] Zhou, X., Tong, W., & Li, L. (2020). Deep learning spatiotemporal air pollution data in China using data fusion. *Earth Science Informatics*, *13*(3), 859-868.