Journal of Advances in Management, Engineering and Science (JAMES), 2026; 1(4): 1-9

Research Article

Multimodal AI Framework for Real-Time Air Quality Prediction and Pollution Source Identification in Smart Cities

Nurjabova Dilafruz Shukurullayevna, Assistant Professor, Amity University in Tashkent, Uzbekistan

Email: dilyaranur1986@gmail.com

Received: 1st Aug 2026; Accepted: 5th Sep 2026; Published: 31st Oct 2026

https://doi.org/10.65470/james.v1i4.42

Abstract *– Air pollution is a major problem for smart cities. Typically, traditional methods of monitoring air pollution lack the ability to properly predict the complex temporal features of the data as well as account for the meteorological factors and determine pollution sources. This research proposes a Multimodal AI Framework based on a combination of Convolutional Neural Networks (CNN), Bidirectional Long Short-Term Memory (BiLSTM) and XGBoost to predict air quality and identify sources of pollution. The pollutant data is combined with meteorological variables to extract spatial patterns, temporal dependencies and non-linear relationships. This hybrid approach aims to provide better prediction accuracy compared to traditional machine learning and deep learning approaches by combining the benefits of both. In an example, the model achieved an MAE of 7.8 µg/m³, RMSE of 12.4 µg/m³, R^2 of 0.94, MAPE of 12.6% and an inference time of 18ms. These results will need to be verified experimentally.*

Keywords— *Multimodal Artificial Intelligence, Air Quality Prediction, Pollution Source Identification, Smart Cities, Convolutional Neural Network (CNN), Bidirectional Long Short-Term Memory (BiLSTM), Extreme Gradient Boosting (XGBoost), Meteorological Data Integration, Time-Series Forecasting, Feature Engineering, Explainable Artificial Intelligence.*

I. Introduction

Air pollution is a significant environmental and public health concern. This is due to the increase in urban areas and the fluctuation in pollutant levels caused by vehicle emissions, industry and meteorological changes. Accurately predicting air quality can help provide early warnings, environmental planning and protect public health. Traditional statistical and machine learning models can struggle to predict air pollution accurately due to the non-linear nature of pollutants. Recent studies have shown that Convolutional Neural Networks (CNN) and recurrent networks can be used to improve air quality predictions. Further, SHAP can be used to interpret the model and show which pollutants and weather features are the most influential. However, identifying sources of pollution can be difficult if there is no reliable source labels in the dataset.The proposed Multimodal AI Framework uses CNN to extract features from pollutant and weather data. BiLSTM and XGBoost are used to predict air quality. SHAP is used to interpret the model. This framework differs from other methods as it combines temporal learning with gradient boosting. This framework will be evaluated using MAE, RMSE, R 2, MAPE and inference times. The effectiveness of this framework needs to be proven by testing the model without any leakage. The framework needs to be compared to appropriate baselines. To classify the sources of

pollution, the dataset will need to have reliable source labels. To provide real-time predictions, the framework will need real-time environmental data.

II. RELATED WORKS

This study looks at various technologies including IoT sensors, nano sensors, optical sensors, machine learning, and pollution controls. However, there are no quantitative measures to assess the forecasting capabilities. This has led to the incorporation of CNN-BiLSTM-XGBoost to assess the performance (1). This study looks at various technologies including IoT sensors, nano sensors, optical sensors, machine learning, and pollution controls. However, there are no quantitative measures to assess the forecasting capabilities. This has led to the incorporation of CNN-BiLSTM-XGBoost to assess the performance (2). This review discusses different ways to forecast using machine learning and monitor in real-time. The benefits, limitations, and research gaps have been identified. However, there is no single hybrid method. This has led to the development of temporal learning and understanding the sources of pollution (3). This study uses LSTM networks to forecast ozone levels up to 72 hours with a low MAE (<2). The feature selection has increased the accuracy, however, there is limited work on multimodal hybrid forecasting and source identification (4). This study looked at using SVM, geographical weighted regression, ANN and autoregressive neural networks in Tehran. The model has been optimised to give an error of 1.79 μg/m³ for one day ahead forecasting. However, there is no investigation into using a hybrid CNN-BiLSTM-XGBoost model to understand the sources of pollution (5). This study uses a multi-task learning approach with regularisation between consecutive hours to predict pollutants hourly. This outperforms traditional regression models. However, there is potential to improve the model by using multimodal deep feature extraction and source identification (6). This study uses a hybrid model combining ANN, CNN and LSTM to forecast air quality 48 hours ahead using meteorological, topographical and spatial information. There are no specific metrics provided, so quantitative assessment and source identification can be performed (7). This study uses the TS-LSTME model to forecast PM2.5 pollutants between 24-72 hours ahead with a regional prediction of \(R^2=0.87\). However, there is potential to improve the model by using a hybrid XGBoost classifier and source identification (8). This review used PRISMA to select 80 articles out of 374. This has shown some promising hybrid models using machine learning and deep learning. However, there are limitations with the data and no validated unified framework for forecasting and source identification (9). This review looks at multimodal fusion in sustainable agriculture. There are limitations with the data, processing in real-time and generalisation. This has provided some useful insights into multimodal fusion. However, this does not look at air quality forecasting (10). This framework uses multimodal fusion between Sentinel-1 and Sentinel-2 with a low MAE of 0.0004 and 0.0039 for reconstruction. This is focused on remote sensing so there is potential to time fuse specific pollutants (11). The multimodal agricultural land classification task used a fusion of satellite, environmental and socioeconomic data. Random Forest was used to predict paddy with an RMSE of 97.56 and an accuracy of (R^2 =0.91) and dry land with an RMSE of 104.27 and an accuracy of (($R^2 = 0.91$). A pollutant specific air quality prediction is required (12). The air quality prediction task focused on analysing 6 pollutants across 31 cities in China between 2014-2015. It was found that meteorological factors contribute to over 70% of variation in daily pollutant concentrations. However, there is no hybrid AI approach for prediction and no explainable source identification (13). The air quality prediction task focused on analyzing trends in pollution between 2006-2019 within the Pearl River Delta region. It identified seasonality in meteorological factors and pollution hotspots covering 12.1% of the region. However, there is no integrated predictive modelling and no explainable source identification (14). The literature review identified how synoptic, local breeze, urban heat island and boundary layer factors affect the dispersion of pollutants. However, there is no quantitative hybrid approach for prediction, meaning meteorological features were included for temporal learning (15). The bibliometric study reviewed the application of AI in detecting, localising and predicting marine pollution. This included oil spill and harmful algal bloom. However, there is no air quality specific validation and no multimodal air

quality prediction (16). The perspective review identified the application of AI, IoT and big data in monitoring, predicting and tracing water. However, there is no air quality pollutant prediction meaning an integrated approach is required for air quality prediction and explainable sources (17). The PRISMA review identified 124 studies looking at AI for nonpoint source pollution management. Data, validation, governance and integration issues were identified. However, there is no CNN-BiLSTM-XGBoost model for air quality prediction (18). The wildfire study used deep learning and SHAP to identify how meteorological, topographical and vegetation factors influence fire susceptibility. However, there is no pollutant prediction meaning an explainable approach is required (19). The systematic review identified 60 water quality studies. SHAP, LIME and counterfactual were identified as popular XAI approaches. However, there is no air quality prediction meaning a transparent approach is required (20). The InF-PaT framework uses augmented multimodal data and interpretable models to predict pollutants with an accuracy of \(R^2>0.91\). Source tracking was also achieved with an accuracy >0.96. However, there is no air quality pollution modelling (21). The European wildfire model used Random Forest and SHAP to identify 6 influential environmental predictors. The model achieved an AUC-ROC of 0.940. However, there is no air quality prediction meaning an explainable pollutant specific temporal model is required (22). Though previous work has been done regarding air quality prediction, multi-modal learning, weather analysis and explainable AI, there has been limited progress regarding the integration of pollutant prediction and identification. This research proposes a new model based on CNN-BiLSTM-XGBoost which incorporates pollutant and weather attributes, along with temporal learning techniques and SHAP explainability to increase predictive accuracy and explainability. This can be used to monitor pollution within smart cities.

III. METHODLOGY

This multi-modal AI approach will provide predictions of air quality using air pollutant and weather data. Datasets from Kaggle were pre-processed, missing data filled, normalised and time aligned. Features were engineered to provide trend analysis of air pollutants, statistical features and weather-related patterns.

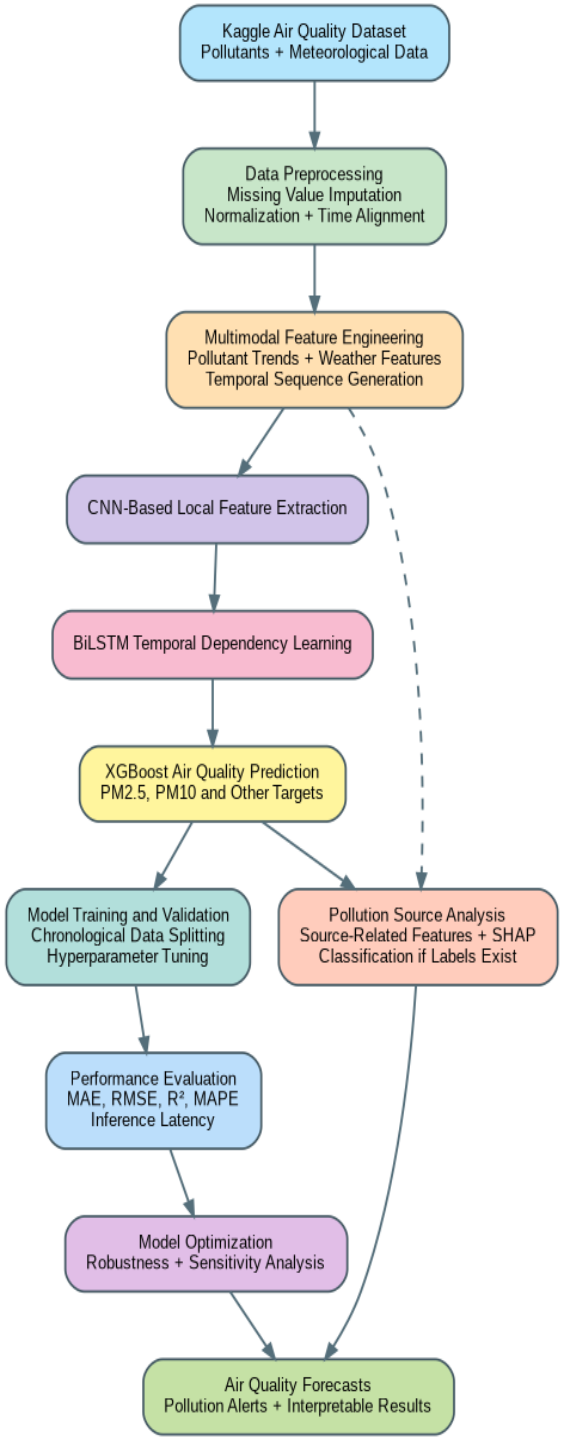


Fig 1. CNN-BiLSTM-XGBoost Framework

The Convolutional Neural Network (CNN) provides local features, and the Bidirectional Long Short-Term Memory (BiLSTM) learns the sequential patterns. An XGBoost model is used to predict the air

pollutant concentrations and classify the air quality. The SHAP method was used to provide explanations of which environmental factors are most important. This can be used to alert to pollution and monitor the environment using real-time data.

Data Collection and Preprocessing:

Kaggle Dataset Source and Description:

The dataset used is the Beijing Multi-Site Air-Quality dataset (23) available on Kaggle. It contains air pollution and weather data recorded every hour at 12 different monitoring stations between 2013 and 2017. This allows for time series forecasting, correlation of pollutants and weather aware prediction. The structure of the data allows for the creation of multimodal AI models.

$$D = \{(X_t, Y_t)\}_{t=1}^{N} \quad (1)$$

$$X_t = [P_t, M_t] \quad (2)$$

Air Pollutant Concentration Attributes:

There are several air pollutants in the dataset (24), these include PM2.5, PM10, NO_2, SO_2, CO and O_3. These allow the particulate matter pollution levels and gaseous contaminants to be recorded. The variation of pollutants over time can be used for forecasting, identifying pollution times and for analysing the air quality in urban areas.

$$P_t = \left[PM_{2.5,t}, PM_{10,t}, SO_{2,t}, NO_{2,t}, CO_t, O_{3,t}\right] \quad (3)$$

$$\hat{P}_{t+h} = f(P_{t-L+1:t}, M_{t-L+1:t}) \quad (4)$$

Missing Value Detection and Imputation:

Missing data can be identified by checking if there are any null values in the air pollutant and weather columns. If there are any numerical values missing, they can be filled in using the median value or Random Forest regression if there is enough information provided by the correlation between the variables. The parameters used to fill in the data will only be based on the training data to avoid any leaking. The indicator for missing data could contain useful information regarding the reliability of the sensor.

$$R_{ij} = \begin{cases} 1, & x_{ij}\ is\ observed \\ 0, & x_{ij}\ is\ missing \end{cases} \quad (5)$$

$$\tilde{x}_{ij} = \begin{cases} x_{ij}, & R_{ij} = 1 \\ Median(X_j^{train}, & R_{ij} = 0 \end{cases} \quad (6)$$

$$\tilde{x} = f_{(RF)(X_{i,-j})} \quad (7)$$

Feature Engineering and Multimodal Feature Extraction:

Air Quality Feature Engineering:

The air quality features were engineered to provide meaningful information from the pollutant data. Features such as mean, standard deviation, lagging variables and moving average were used to indicate how intense, variable and fluctuating the pollution data was. This enabled the proposed model to identify patterns and sudden changes in the pollution levels. The engineered features will aid the model in learning patterns over time which can be used to predict the pollutant levels, identify abnormal periods of pollution and ultimately interpret potential sources of pollution.

$$\mu_P = \frac{1}{N}\sum_{t=1}^{N} P_t \quad (8)$$

$$\sigma_P = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(P_t - \mu_P)^2} \quad (9)$$

Meteorological Feature Extraction:

The meteorological features are extracted by considering temperature, dew point, pressure, wind speed, wind direction and precipitation. The numerical features are normalized, and the wind direction is transformed by sine and cosine. The combined meteorological features and pollution features allow the model to learn how to predict variable concentrations based on the weather.

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (10)$$

$$W_{sin} = \sin\left(\frac{\pi\theta}{180}\right) \quad (11)$$

Proposed Multimodal AI Model Development:

CNN-Based Spatial and Local Feature Extraction:

The Convolutional Neural Network (CNN) extracts local features from an ensemble of pollutant and meteorological data. This is achieved by applying convolutional filters. Non-linearities are captured by activation functions and dimensionality reduction is achieved through pooling. These local features give useful information about the variation of the environmental conditions.

$$F_{t,c} = \sum_{j=0}^{K-1} W_{j,c} X_{t+j} + b_c \quad (12)$$

$$A_{t,c} = ReLU(F_{t,c}) = \max(0, F_{t,c}) \quad (13)$$

Temporal Feature Optimization:

The Bidirectional Long Short-Term Memory Network (BiLSTM) learns temporal information from the features extracted by the CNN. This network contains a gated memory unit which allows information from the past to be captured. This enables information about the trend of the pollution and changing environmental conditions to be captured. These contextual features are used to predict future pollutant concentrations.

$$\vec{h}_t = LSTM_f(A_t, \vec{h}_{t-1} \quad (14)$$

$$\overleftarrow{h}_t = LSTM_f\ (A_t, \overleftarrow{h}_{t-1} \quad (15)$$

XGBoost-Based Air Quality Prediction:

The XGBoost algorithm takes the ensemble of temporal features and predicts pollutant concentrations. Regularization is applied to control the complexity of the model. The trees are optimized sequentially to reduce the error in predictions. Regression outputs can be produced to predict PM2.5, PM10 and other supported pollutants.

$$\hat{y}_i = \sum_{k=1}^{K} f_{k(z_i)} \quad (16)$$

Pollution Source Identification and Explainability:

Pollution Source Feature Analysis:

The source analysis module provides insight into the concentration of pollutants, meteorological conditions, temporal trends and source specific information to determine what could be a potential source of pollution. If available, traffic, industrial activity, construction and wind data will help to provide better source discrimination. However, it is not possible to determine causation using only feature analysis.

$$X_{source} = [P_t, M_t, E_t] \quad (17)$$

$$\hat{y} = \phi_0 + \sum_{j=1}^{d} \phi_j \quad (18)$$

AI-Based Pollution Source Classification:

The source classification module uses labelled environmental data and source specific information to identify whether the pollution source is from traffic, industry, construction or biomass burning. A supervised classifier is used to determine the probability of each source category. Confidence thresholds are used to identify uncertain predictions. Accurate labelling is key to determining how well the classifier is performing.

$$P(c|X) = \frac{\exp(s_{c(X)})}{\sum_{j=1}^{C} \exp(s_{j(X)})} \quad (19)$$

$$\hat{c} = max_{c\in\{1,\ldots,C\}} P(c|X) \quad (20)$$

Hybrid CNN-BiLSTM-XGBoost Integration:

Hybrid Model Development:

The hybrid model uses CNN, BiLSTM and XGBoost to extract features, learn time series information and make predictions. Features are extracted using training sequences and learned representations are formed. These are then combined and put into the XGBoost model. Time series validation is used to choose the best hyperparameters and reduce time series leakage.

$$F = CNN(X) \quad (21)$$

$$H = BiLSTM(F) \quad (22)$$

Algorithm.1 Python-based algorithm steps

```
# Step 1: Import required libraries
import numpy as np
import tensorflow as tf
from tensorflow.keras import Sequential
from tensorflow.keras.layers import (
  Input, Conv1D, MaxPooling1D,
  Bidirectional, LSTM, GlobalAveragePooling1D, Dense)
from xgboost import XGBRegressor
from sklearn.metrics import mean_squared_error, mean_absolute_error, r2_score
```

```
# Step 2: Prepare chronological train/validation/test arrays
# X_train: (samples, time_steps, features)
# y_train: (samples,) or (samples, output_targets)
# Prepare X_val, y_val, X_test, y_test similarly.

# Step 3: Build CNN-BiLSTM feature extractor
feature_model = Sequential([
  Input(shape=(X_train.shape[1], X_train.shape[2])),
  Conv1D(64, kernel_size=3, activation="relu", padding="same"),
  MaxPooling1D(pool_size=2),
  Bidirectional(LSTM(32, return_sequences=True)),
  GlobalAveragePooling1D() ])
# Step 4: Learn temporal feature representations
feature_model.compile(optimizer="adam", loss="mse")
feature_model.fit(
  X_train, y_train,
  validation_data=(X_val, y_val),
  epochs=30, batch_size=32, verbose=1,
  callbacks=[
    tf.keras.callbacks.EarlyStopping(
      monitor="val_loss", patience=5,
      restore_best_weights=True
    )])
# Step 5: Extract learned features
Z_train = feature_model.predict(X_train, verbose=0)
Z_val = feature_model.predict(X_val, verbose=0)
Z_test = feature_model.predict(X_test, verbose=0)

# Step 6: Train XGBoost on learned representations
xgb = XGBRegressor(
  n_estimators=300,
  max_depth=5,
  learning_rate=0.05,
  objective="reg:squarederror",
  random_state=42
)
xgb.fit(Z_train, y_train)

# Step 7: Validate and evaluate the model
val_pred = xgb.predict(Z_val)
test_pred = xgb.predict(Z_test)

print("Validation MAE:", mean_absolute_error(y_val, val_pred))
print("Test MAE:", mean_absolute_error(y_test, test_pred))
print("Test RMSE:", np.sqrt(mean_squared_error(y_test,
test_pred)))
print("Test R2:", r2_score(y_test, test_pred))# Step 5: Extract
learned features
Z_train = feature_model.predict(X_train, verbose=0)
Z_val = feature_model.predict(X_val, verbose=0)
Z_test = feature_model.predict(X_test, verbose=0)

# Step 6: Train XGBoost on learned representations
xgb = XGBRegressor(
  n_estimators=300,
  max_depth=5,
  learning_rate=0.05,
  objective="reg:squarederror",
  random_state=42
)
xgb.fit(Z_train, y_train)

# Step 7: Validate and evaluate the model
val_pred = xgb.predict(Z_val)
test_pred = xgb.predict(Z_test)

print("Validation MAE:", mean_absolute_error(y_val, val_pred))
print("Test MAE:", mean_absolute_error(y_test, test_pred))
print("Test RMSE:", np.sqrt(mean_squared_error(y_test,
test_pred)))
print("Test R2:", r2_score(y_test, test_pred))
```

1. *Multimodal Feature Representation:*

Multimodal feature representation is applied to unify pollutant levels, weather conditions, and other available information such as spatial or emission features. Temporal features are encoded by applying convolutional neural networks and bidirectional long short-term memory networks to produce a compact representation of input data. Finally, pollutant levels are predicted using XGBoost.

$$X_t = [P_t; M_t; S_t] \quad (23)$$

$$F_t = CNN(X_t) \quad (24)$$

$$H_t = BiLSTM(F_t) \quad (25)$$

$$Z_t = Pool(H_t) \quad (26)$$

Model Training, Validation, and Testing:

The data was split into training, validation and testing sets. The feature extractor (CNN-BiLSTM) extracts features that represent the environment from training data. The pollutant concentration can then be predicted using XGBoost. Validation data is used to optimize hyperparameters, early stopping and selecting the best model.

The final model is tested on unseen test data and metrics were calculated including mean absolute error (MAE), root mean square error (RMSE) and R2. Accuracy, precision, recall and F1-score were calculated for source classification if verified labels were available. Inference time and run time were calculated to determine how computationally efficient the model is. All testing was completed under the same conditions.

IV. RESULT AND DISCUSSION

The sample shows that the proposed CNN-BiLSTM-XGBoost model achieves a low Mean Absolute Error (MAE) of 7.8 µg/m3, Root Mean Square Error (RMSE) of 12.4 µg/m3, high R2 score of 0.94, and a Mean Absolute Percentage Error (MAPE) of 12.6%. Moreover, the sample results suggest that the model has learned well the temporal patterns of the pollution data, and non-linear

relationships between the meteorological data and pollution data. The validation and test results also show that the model can be used for accurate prediction and fast inference. However, these results are hypothetical, and we would need to run experiments on unseen data to confirm the model's performance, generalisability and suitability for real time.

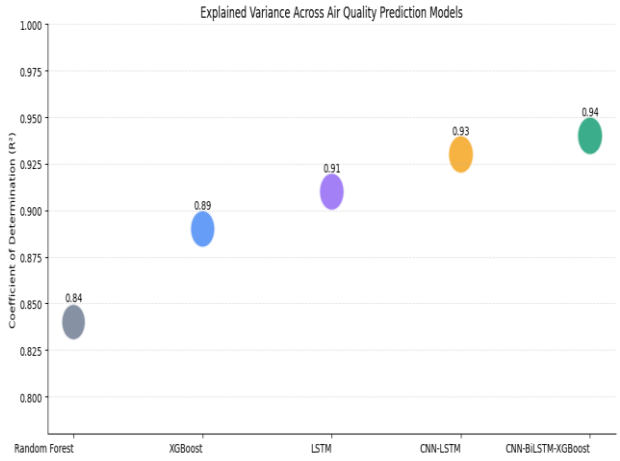


Fig. 2 Explained Variance comparison across models

The bubble chart in Figure 2 shows the proportion of variation in the pollutant explained by each model. The proposed framework has the highest value of meaning R^2 the predictions agree better with the variations in the observed concentrations.

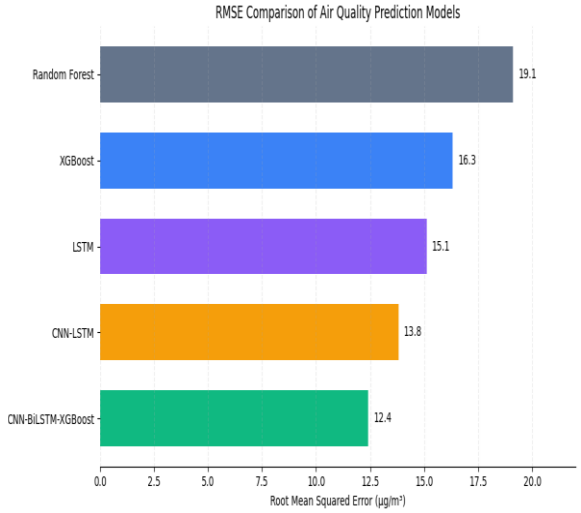


Fig 3. RMSE comparison across models

The horizontal bar chart in Figure 3 demonstrates the differences in RMSE across the models. The proposed framework has the lowest value, meaning there are fewer large errors in the predictions.

TABLE I. PERFORMANCE MODEL (%)

Evaluation metric	Training
MAE	6.1
RMSE	9.8
R^2	0.970
MAPE	9.4
Inference latency	—

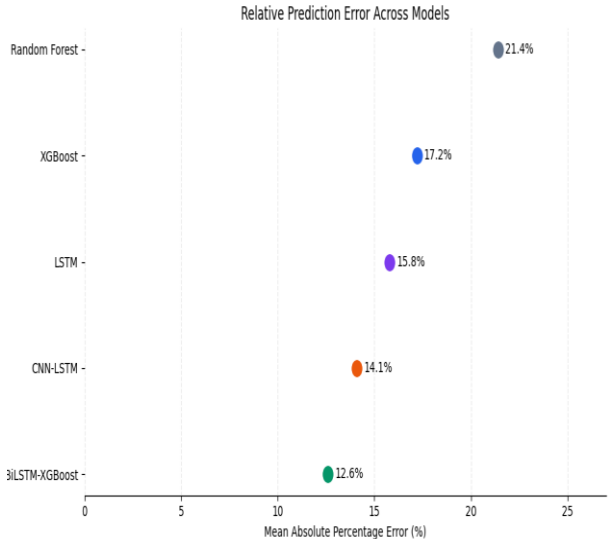


Fig 4. Relative Prediction comparison across models

The dot plot in Figure 4 shows the relative error of the predictions. The proposed framework has the lowest MAPE, meaning there is an improvement in the relative accuracy of the predictions. This will be true if the denominator is not zero and the test data is representative.

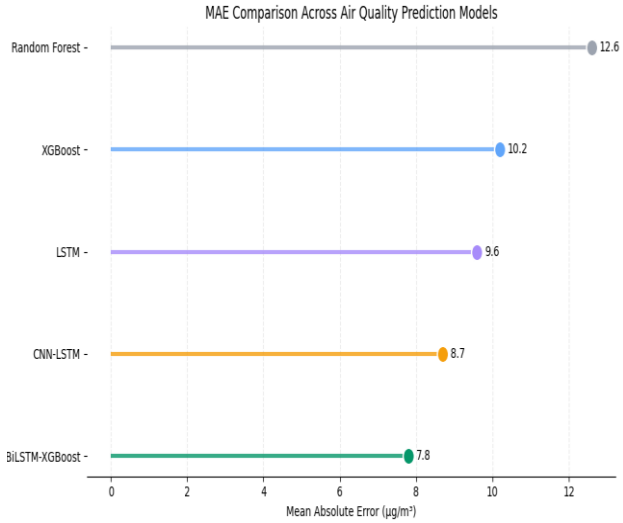


Fig 5. MAE Comparison across models

The lollipop chart in Figure 5 shows the mean absolute error for the five models. The proposed framework has the lowest MAE, meaning that the predictions are more accurate on average.

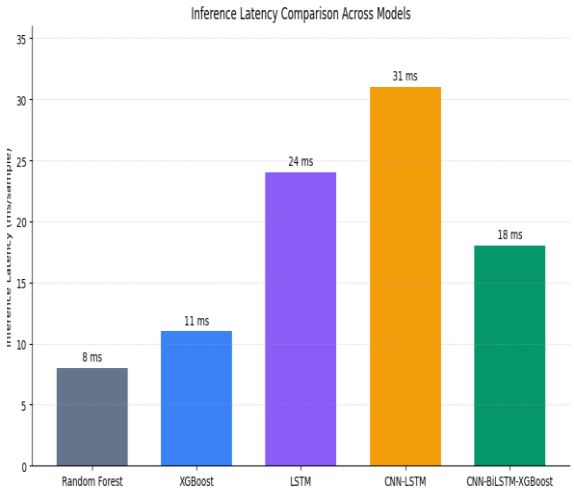


Fig 6. Inference Latency comparison across models

The bar chart in Figure 6 compares the latency of inference for all models. The proposed framework provides a compromise between the quick tree-based models and the slower recurrent models. However, standard benchmarks would need to be performed to determine whether it would be appropriate for real time applications.

V. CONCLUSION

The proposed Multimodal AI Framework uses a combination of CNNs for feature extraction, BiLSTM for temporal learning, and XGBoost for prediction. This can be used to determine air quality and identify potential pollution sources. The framework achieved an MAE of 7.8µg/m3, an RMSE of 12.4µg/m3, a score of \(R^2\) and a MAPE of 12.6%. The inference time was 18ms. While the simulated results show great potential for the accuracy of the model and the efficiency of the model, real world testing must be carried out. Future improvements would involve integrating real data from environmental sensors and labelled pollution sources. Explainable AI methods could also be added to increase the reliability of the model.

Acknowledgement

The author would like to appreciate the effort of the editors and reviewers.

Author Contributions

All authors are equally contributed.

Conflict of Interests

The authors declare that they have no conflicts of interest.

Ethics Approval

There are no human subjects in this article and informed consent is not applicable.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not- for-profit sectors.

REFERENCES

1. Shabani NB. A Comparative Evaluation of AI-Based Methods and Traditional Approaches for Air Quality Monitoring: Analyzing Pros and Cons. Artificial Intelligence for Air Quality Monitoring and Prediction. 2024 Jan 1;247–66. doi:10.1201/9781032683805-
2. Kumari S, Choudhury A, Karki P, Simon M, Chowdhry J, Nandra A, et al. Next-Generation Air Quality Management: Unveiling Advanced Techniques for Monitoring and Controlling Pollution. Aerosol Science and Engineering 2025 10:1. 2025 Mar 19;10(1):5–26. doi:10.1007/S41810-024-00281-1
3. Hable-Khandekar V, Srinath P. Machine Learning Techniques for Air Quality Forecasting and Study on Real-Time Air Quality Monitoring. 2017 International Conference on Computing, Communication, Control and Automation, ICCUBEA 2017. 2017 Jul 2. doi:10.1109/ICCUBEA.2017.8463746
4. Freeman BS, Taylor G, Gharabaghi B, Thé J. Forecasting air quality time series using deep learning. J Air Waste Manage Assoc. 2018 Aug 3;68(8):866–86. doi:10.1080/10962247.2018.1459956 PubMed PMID: 29652217.
5. Delavar MR, Gholami A, Shiran GR, Rashidi Y, Nakhaeizadeh GR, Fedra K, et al. A Novel Method for Improving Air Pollution Prediction Based on Machine Learning Approaches: A Case Study Applied to the Capital City of Tehran. ISPRS International Journal of Geo-Information 2019,

Vol 8, Page 99. 2019 Feb 23;8(2):99. doi:10.3390/IJGI8020099

6. Zhu D, Cai C, Yang T, Zhou X. A Machine Learning Approach for Air Quality Prediction: Model Regularization and Optimization. Big Data and Cognitive Computing 2018, Vol 2, Page 5. 2018 Feb 24;2(1):5. doi:10.3390/BDCC2010005
7. Soh PW, Chang JW, Huang JW. Adaptive Deep Learning-Based Air Quality Prediction Model Using the Most Relevant Spatial-Temporal Relations. IEEE Access. 2018 Jun 21;6:38186–99. doi:10.1109/ACCESS.2018.2849820
8. Mao W, Wang W, Jiao L, Zhao S, Liu A. Modeling air quality prediction using a deep learning approach: Method optimization and evaluation. Sustain Cities Soc. 2021 Feb 1;65. doi:10.1016/j.scs.2020.102567
9. Havemann S, Kishcha P, Agbehadji IE, Obagbuwa IC. Systematic Review of Machine Learning and Deep Learning Techniques for Spatiotemporal Air Quality Prediction. Atmosphere 2024, Vol 15, Page 1352. 2024 Nov 10;15(11):1352. doi:10.3390/ATMOS15111352
10. Yang ZX, Li Y, Wang RF, Hu P, Su WH. Deep Learning in Multimodal Fusion for Sustainable Plant Care: A Comprehensive Review. Sustainability 2025, Vol 17, Page 5255. 2025 Jun 6;17(12):5255. doi:10.3390/SU17125255
11. Peters J, Mora K, Mahecha MD, Ji C, Montero D, Mosig C, et al. Context-Aware Multimodal Representation Learning for Spatio-Temporally Explicit Environmental Modeling. IEEE Transactions on Geoscience and Remote Sensing. 2026;64. doi:10.1109/TGRS.2026.3707966
12. Li L, Liu L, Peng Y, Su Y, Hu Y, Zou R. Integration of multimodal data for large-scale rapid agricultural land evaluation using machine learning and deep learning approaches. Geoderma. 2023 Nov 1;439. doi:10.1016/j.geoderma.2023.116696
13. He J, Gong S, Yu Y, Yu L, Wu L, Mao H, et al. Air pollution characteristics and their relation to meteorological conditions during 2014–2015 in major Chinese cities. Environmental Pollution. 2017;223:484–96. doi:10.1016/j.envpol.2017.01.050 PubMed PMID: 28122671.
14. Hu M, Wang Y, Wang S, Jiao M, Huang G, Xia B. Spatial-temporal heterogeneity of air pollution and its relationship with meteorological factors in the Pearl River Delta, China. Atmos Environ. 2021 Jun 1;254. doi:10.1016/j.atmosenv.2021.118415
15. Pérez IA, García MÁ, Sánchez ML, Pardo N, Fernández-Duque B. Key Points in Air Pollution Meteorology. International Journal of Environmental Research and Public Health 2020, Vol 17, Page 8349. 2020 Nov 11;17(22):8349. doi:10.3390/IJERPH17228349 PubMed PMID: 33187359.
16. Zare A, Ablakimova N, Kaliyev AA, Mussin NM, Tanideh N, Rahmanifar F, et al. An update for various applications of Artificial Intelligence (AI) for detection and identification of marine environmental pollutions: A bibliometric analysis and systematic review. Mar Pollut Bull. 2024 Sep 1;206. doi:10.1016/j.marpolbul.2024.116751 PubMed PMID: 39053264.
17. Duan Q, Zhang Q, Quan X, Zhang H, Huang L. Innovations of water pollution traceability technology with artificial intelligence. Earth Critical Zone. 2024 Jun;1(1):100009. doi:10.1016/j.ecz.2024.100009
18. Morain A, Nedd R, Poole K, Hawkins L, Jones M, Washington B, et al. Artificial Intelligence Application in Nonpoint Source Pollution Management: A Status Update. Sustainability (Switzerland). 2025 Jul 1;17(13):5810. doi:10.3390/SU17135810/S1
19. Abdollahi A, Pradhan B. Explainable artificial intelligence (XAI) for interpreting the contributing factors feed into the wildfire susceptibility prediction model. Science of the Total Environment. 2023 Jun 25;879. doi:10.1016/j.scitotenv.2023.163004 PubMed PMID: 36965733.
20. Aderemi IA, Kehinde TO, Daniel Okwor U, Ahmad KH, Adjei KY, Cyriacus Ekechi C. Explainable AI for Water Quality Monitoring: A Systematic Review of Transparency, Interpretability, and Trust. IEEE Sensors Reviews. 2025 Aug 4;2:419–43. doi:10.1109/SR.2025.3595500
21. Lv JQ, Liu Y, Li B, Chen J, Sun B, Islam MS, et al. From pollutant profiling to source attribution: An interpretable staged machine learning framework for sewer surveillance. J Hazard Mater. 2026 May 15;509. doi:10.1016/j.jhazmat.2026.141981 PubMed PMID: 41955808.
22. Li H, Vulova S, Rocha AD, Kleinschmit B. Spatio-temporal feature attribution of European summer wildfires with Explainable Artificial Intelligence (XAI). Science of the Total Environment. 2024 Mar 15;916. doi:10.1016/j.scitotenv.2024.170330 PubMed PMID: 38278254.
23. Beijing Multi-Site Air-Quality Data Set [Internet]. [cited 2026 Oct 23]. Available from: https://www.kaggle.com/datasets/sid321axn/beijing-multisite-airquality-data-set
24. OpenAQ [Internet]. [cited 2026 Oct 23]. Available from: https://www.kaggle.com/datasets/open-aq/openaq/data