Nature Environment and Pollution Technology
An International Quarterly Scientific Journal

p-ISSN: 0972-6268
(Print copies up to 2016)
e-ISSN: 2395-3454

Vol. 25 | No. 3 | Article ID B4381 | 2026

Original Research Paper | https://doi.org/10.46488/NEPT.2026.v25i03.B4381 | Open Access Journal

Deep Learning for AQI Prediction Using Multiple Feature Vectors: A Case Study of Colaba and Deonar Stations

Darakhshan Khan[1]†, Archana B. Patankar[1], Himani Deshpande[2] and Juhi Janjua[1]

[1]Department of Computer Engineering, Thadomal Shahani Engineering College, Mumbai, 400050, India
[2]Department of Artificial Intelligence and Data Science, Thadomal Shahani Engineering College, Mumbai, 400050, India
†Corresponding author: Darakhshan Khan; darakshan.khan@thadomal.org

Abbreviation: Nat. Env. & Poll. Technol.
Website: www.neptjournal.com

Received: 26-07-2025
Revised: 29-09-2025
Accepted: 07-10-2025

Key Words:

Air pollutants
Meteorological parameters
Bidirectional LSTM
Deep learning for AQI prediction

Citation for the Paper:

Khan, D., Patankar, A.B., Deshpande, H. and Janjua, J., 2026. Deep learning for AQI prediction using multiple feature vectors: A case study of Colaba and Deonar stations. *Nature Environment and Pollution Technology*, 25(3), B4381. https://doi.org/10.46488/NEPT.2026.v25i03.B4381

Note: From 2025, the journal has adopted the use of Article IDs in citations instead of traditional consecutive page numbers. Each article is now given individual page ranges starting from page 1.

ABSTRACT

Air quality monitoring and prediction are important for effective public health strategies, as air pollution is a major contributor to mortality. The goal is to design and evaluate different deep learning models: Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), and a hybrid Conv1D-LSTM on different input sets of features. The dataset consists of pollutant and meteorological parameters from 2019 to 2024 with hourly frequency for two monitoring stations: Colaba and Deonar. The models were trained on three sets of features: pollutant-only, meteorological-only, and combined features. Model accuracy was determined using the root mean square error, coefficient of determination, mean absolute percentage error, and explained variance score. The results indicate that the combined attributes significantly enhanced the prediction quality, with the hybrid CNN-LSTM performing best at Colaba and the LSTM on meteorological attributes performing best at Deonar. The Bi-LSTM exhibited consistent performance on the feature sets. These results underscore the importance of using both pollutants and meteorological information and illustrate the efficiency of sophisticated deep learning structures for location-based air quality prediction.

1. INTRODUCTION

Globally, air pollution has emerged as a major cause of mortality and a significant threat to human health. Monitoring, understanding, and predicting environmental factors, specifically those related to air quality, remain critical challenges for governments and corporations worldwide. Given the increasing rate of pollution, there is a critical need to develop methods that can accurately predict the Air Quality Index and support the creation of better public health strategies (Schürholz et al. 2020). Human life is basically surviving on air, but its quality gets ruined due to vast pollution, which leads to acute multiple health hazards, such as physiological and respiratory disorders. Scientific studies have proven that air pollution is the greatest environmental risk, driven by rapid industrialization and urban growth. This unplanned emission has grossly degraded air quality, exposed people to hazardous substances, and posed substantial risks to public health.

The need to mitigate these impacts underlines the importance of robust monitoring systems and predictive models in safeguarding air quality and reducing the adverse health effects of pollution on communities (Gupta et al. 2023). Particulate matter (PM) has become a major pollutant of concern, significantly affecting air quality across borders. The particles in PM vary in size and chemical composition; therefore, the risks associated with them are also variable in space and time. Studies have identified a relationship between PM exposure and daily respiratory mortality in both construction workers and residents living near construction sites (Yan et al. 2023), and it can exacerbate chronic obstructive pulmonary disease (COPD) symptoms, thereby increasing the risk of mortality (Gou et al. 2023, DeVries et al. 2017). Other common air pollutants include ammonia, sulfur dioxide, carbon monoxide, nitrogen oxides, and ozone, all of which have adverse effects on human health. Ammonia is an irritating gas that affects the skin, respiratory, digestive, and ophthalmic systems. Ammonia combines with other air constituents to form $PM_{2.5}$, a key air pollutant with enormous health effects (Petrus et al. 2022). Both SO_2 and ozone have been shown to significantly increase the risk of hospital admission, cardiopulmonary diseases, myocardial infarction, chronic obstructive pulmonary disease, and respiratory diseases (Wu et al. 2022, Jian et al. 2024). Higher levels of these pollutants are directly correlated with the proportion of the population with chronic respiratory diseases. Increased levels of CO in the air can lead to a reduction in oxygen levels in the human body, which can damage cells, tissues, and organs, whereas high NOx levels can cause respiratory morbidity and lung diseases (Zhang & Zhang 2024).

Climate change is evident across the globe, and variations in climatic conditions have a strong influence on the Air Quality Index (AQI) levels, which, if not controlled, can amplify risks to human health (Liu et al. 2022, Yan et al. 2024). Multiple studies have shown a strong correlation between AQI levels and various climatic factors, such as temperature, wind direction, wind speed, humidity, pressure, visibility, rainfall, sun exposure, diurnal temperature, precipitation, and weather outlook conditions. Multiple studies have claimed that the principal climatic factors affecting AQI levels are temperature, pressure, humidity, dew point, and wind speed. Owing to local geomorphological and meteorological factors, AQI levels are affected within a radius of 100 km (Liu et al. 2022).

With the advancement of deep learning, various sequential deep architectures, such as the Recurrent Neural Network (RNN), Gated Recurrent Unit (GRU), and Long Short-Term Memory cell (LSTM), have gained popularity in making accurate and robust AQI predictions for many cities worldwide. An RNN is created from a feed-forward neural network, wherein some nodes have inter-node connections and backward loops that allow the network to remember the present and recent past (Ayus et al. 2023). A variant of RNN, the GRU, utilizes two main gates, an update gate and a reset gate, to manage memory. The update gate controls the addition of new information, and the reset gate controls the amount of the previous state to remember. It can be further extended as a Bidirectional GRU, which learns sequences in both directions by combining two unidirectional GRUs for learning in the forward and backward directions (Ayus et al. 2023, Mandal & Sen 2024). Another variant of RNN is LSTM, which provides a solution to the problems of long-term dependency with forget, input, and output gates that control the memorization and removal of information (Ayus et al. 2023, Song et al. 2024). The Bidirectional LSTM further improves this by processing input sequences in both forward and backward directions. CNN-GRU is a hybrid model in which a Convolutional Neural Network is applied for feature extraction and a GRU for sequence prediction, thus creating a deep architecture capable of handling spatiotemporal data (Ayus et al. 2023, Shi et al. 2023).

This study involved performing a statistical correlation analysis between the AQI and various factors, such as meteorological parameters and atmospheric pollutants. It also aims to understand the effects of diurnal and seasonal variations on air quality data. The core goal is to understand how various feature vectors interact simultaneously with different deep learning architectures. In addition, this study seeks to determine which deep learning architectures provide the most robust predictions and have better generalization abilities. Two stations with different environments were studied: Colaba, which is located on the coast, and Deonar, an inland station influenced mainly by landfill and industrial pollution sources.

The remainder of this paper is organized as follows: the next section contains a review of the literature, which highlights the global perception of AQI prediction, advancements in machine learning and deep learning techniques, and the role of climatic conditions in AQI forecasting. The following section encompasses an understanding of the relationship between AQI and the factors influencing it, followed by temporal and spatial analysis. The following section deals with the adopted methodology, including detailed information on the data collection process, different model architectures used, feature engineering, and model validation. The final section presents the setup of the experiment, results, and observations of the research, followed by the conclusion.

2. BACKGROUND RESEARCH

Air pollution has emerged as a global health crisis; therefore, stable and steady forecasting models for air quality monitoring and prediction are necessary. Different researchers have utilized different methodologies to solve this problem, especially through enhanced machine learning and deep learning techniques. This section discusses some of the most significant studies worldwide to provide a complete context for the current research.

2.1 Cross-Continental Perspective on AQI Prediction

China: Exhaustive research in China has shown that variations in AQI levels are largely influenced by atmospheric pollutants such as PM2.5, PM10, Ozone (O_3), SO_2, NOx, and CO, and the concentration of these atmospheric pollutants is heavily influenced by climatic conditions. The studies (Zhang & Jiang 2018, Sun et al. 2019) use positive matrix factorization with chemical mass balance, regression, and correlation analysis on data collected from different monitoring stations across multiple provinces in China, to gauge the association between the concentration of PM2.5 and various meteorological factors, such as temperature, wind speed, relative humidity, and air pressure. It was observed that low wind speed, coupled with high rising temperature and high humidity, will lead to higher concentrations of PM2.5 in the air. Higher PM2.5 levels were observed in winter and during peak hours of the day. Tian et al. (2021) used wavelet transform and demonstrated that the ozone concentration shows strong temporal and spatial associations with meteorological variables, such as wind speed, solar radiation, temperature, and humidity, with a noteworthy observation that urban areas are prone to higher ozone levels. Whereas another study, in Linfen city, which is one of the most polluted areas in China, seasonal variation analysis was carried out to capture the climatic fluctuations on air pollution. They found that during the winter season, air quality was very poor, which was heavily influenced by the speed of the wind and precipitation (Cui et al. 2018). Fang et al. (2015) performed extensive research on a dataset collected from air quality monitoring stations for 338 cities across China, using multiple tools such as ordinary least squares, special lag models, geographically weighted regression, and Spearman's rank correlation to understand the effect of population size, industrialization, and transportation networks on air pollution.

India: Multiple studies conducted across different cities in India indicate that AQI levels are influenced by meteorological conditions and location-specific factors, such as distance from dumping yards or proximity of monitoring stations to the sea coast. The review article (Karthick et al. 2024) integrates multiple pieces of research conducted in different cities, namely Ahmedabad, Delhi, Lucknow, Gurugram, and Mumbai. This highlights the observation that the concentrations of PM2.5 and CO are the two most crucial factors influencing the AQI; thus, the reduction in the two aforementioned pollutants can contribute significantly to better air quality in Indian cities. Another study (Varaprasad et al. 2024) used five coastal Indian cities: Kolkata, Visakhapatnam, Chennai, Mumbai, and Thiruvananthapuram as the research objects. The results indicate that the highest concentrations of PM2.5 occur during winter (December to February), whereas the lowest concentrations are observed during the monsoon (June to September). Among all these findings, an important outcome is that the eastern Indian coastline bears an intense influence from outflowing emissions from the Indo-Gangetic Plain that gradually decreases from north to south, that is, from Kolkata to Chennai. The analysis of the sea-land breeze cycles indicated a negative correlation between breeze magnitude and PM2.5, suggesting that weaker breezes might lead to worse air quality because of weaker ventilation. The researchers (Rangaswamy et al. 2022, Poyyamoli & Boss 2014) have used data from Jawahar Nagar dump yard, Hyderabad, and Kammiyampet dump yard, Cuddalore, to understand how improper garbage disposal practices around the landfills can increase the concentration of air pollutants in nearby areas. It was observed that due to the decomposition of waste or burning of garbage at a dumping site, there was a much higher concentration of multiple air pollutants than the permissible limits, which can lead to various disorders related to the respiratory system or cardiovascular problems. There was also the formation of a chemical substance called leachate, which pollutes land and groundwater.

America and Europe: Studies from the American and European continents have investigated the interplay between air quality and various human activities, such as urbanization, transportation, and agriculture. Sterling (2024) investigated how air quality varies with population density in any country, as well as the number of transportation-related emissions connected to it, in California. Notable findings are that a higher density is associated with increased emissions of vehicles; thus, the more a region contains human populations, the higher the emissions, meaning more air pollutants from transport. In Romania, where 14% of the civil workforce is engaged in agriculture, a study (Petrus et al. 2022) was conducted to determine the proportion of NH3 using laser photoacoustic spectroscopy. The key sources of higher NH3 levels are agricultural practices and emissions caused by fuel combustion. Other pertinent variables affecting the over-time dispersion and accumulation of NH3 include the temperature, humidity, and wind speed. This study (Sajjad Abdollahpour et al. 2024) illustrates considerable links between urban spatial structure and air quality using a multi-decade analysis of 481 U.S. cities from 1990 to 2015. Compact urban forms, population density, circularity, and green spaces diminish pollutant concentrations, whereas expansion and industrial areas tend to enhance them. A 10% improvement in key factors could potentially prevent over 10,000 deaths annually.

Asia and Africa: In Adıyaman (Kara et al. 2024), the authors claimed that natural and human activities drive air pollution. Major contributors include dust transport from southerly winds, crop burning, and heating, which are primarily found in the city center. PM10 was positively correlated with SO_2, whereas wind speed and temperature had negative effects on pollution levels. Saadi et al. (2021) used a Polar 810i monitor to test CO, Heart Rate Variability (HRV), and city size in Tel Aviv (metropolitan) and Afula (small city). The results showed that pollution is high in Tel Aviv due to industrialization and vehicle emissions, with CO levels showing a positive correlation and HRV showing a negative association, indicating greater cardiac effects. Weak positive correlations between PM2.5 and temperature (r = 0.42) and humidity (r = 0.37) have been reported in a study conducted in Ratchaburi, Thailand. AirQ+ [33] states that PM2.5 during dry periods violated the limit set by WHO, and the elderly with more years lost due to disability. In Addis Ababa, deteriorating vehicle conditions and poor road status are drivers of air pollution, especially in the heavily congested areas of Megenagna. SO_2, NO_2, PM2.5, and PM10 data from 43 sites showed significant spatial heterogeneity, with SO_2 often above the safety thresholds. This study (Bizualem et al. 2023) calls for further research on traffic-related emissions to reduce pollution. Global studies have shown that meteorological factors, urbanization, transportation, and waste management have a significant impact on air quality, with most pollutants, such as PM2.5, PM10, CO, and SO_2, exceeding the safety limit due to human activities and climatic conditions.

2.2 Innovation in AQI Prediction Techniques

Neural network-based Models: These models have garnered significant attention for predicting air quality indices (AQI) and pollutant concentrations. For example, one study that utilized data from four Chinese cities implemented machine learning algorithms such as extreme gradient boosting (XGBoost), light gradient boosting, and random forest. Among these, XGBoost demonstrated the highest forecasting accuracy, achieving a coefficient of determination of 0.929 (Wang et al. 2023). Another comparative analysis (Zhan et al. 2020) explored various predictive models, including linear regression, CNN, LSTM, GRU, Bidirectional LSTM, and backpropagation neural networks. Using hourly AQI data from 1,615 locations across China, this study identified the Bidirectional LSTM as the most effective model.

A broader review (Mendez et al. 2023) examined global approaches to AQI prediction, with a specific focus on PM2.5. It highlighted the frequent use of weather data and pollutant concentrations as input features and identified models such as LSTM, MLP, CNN, and GRU as the most commonly applied techniques. Evaluation metrics such as RMSE, R2 Score, and MAPE were consistently used to assess performance. Additionally, research involving data from the Belfast city center compared the predictive capabilities of LSTM, GRU, and ARIMA models for pollutants, including NO_2, O_3, SO_2, PM2.5, and PM10. The study found that LSTM consistently outperformed other models, whereas ARIMA delivered subpar predictions (Naz et al. 2023). Collectively, these findings underscore the increasing reliance on neural network architectures for precise air quality modeling.

2.3 Hybrid Models for Optimized Accuracy

These approaches to AQI prediction have demonstrated notable improvements by integrating simple and advanced methods. For instance, one study (Kleingchuay et al. 2023) employed empirical mode decomposition to extract time-frequency information, which was combined with Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) models to analyze residuals and historical data. The output was subsequently processed through the LSTM and SVM models, yielding enhanced MAE scores for LSTM and superior index of agreement values for SVM. Another study (Barve et al. 2020), using meteorological data and AQI values from Beijing, introduced a hybrid framework with parallel dense neural networks and LSTM layers. This approach achieved significantly better MAE values than the standalone LSTM models. Similarly, Thakur et al. (2023) combined principal component analysis (PCA) with deep learning models, such as LSTM, GRU, and Bidirectional LSTM, for multivariate AQI forecasting, with PCA and GRU delivering the best performance across various metrics.

For more advanced architectures, Shi et al. (2023) proposed a hybrid TCNbiGRU model that utilized Temporal Convolutional Networks (TCN) to capture long-term patterns and biGRU for short-term dependencies, outperforming standalone LSTM and GRU approaches. In another study (Lu & Li 2023), researchers used a hybrid CNNBiLSTM model optimized through Bayesian methods to predict air pollutants in Tianjin, while another study (Nikpour P et al. 2024) introduced the ”Gelato” framework. This innovative approach combined particle swarm optimization, a transformer-inspired architecture, and XGBoost to predict the multivariate AQI. In research focusing on Indian metropolitan areas (Binbusayyis et al. 2024), a stacked attention-based GRU model incorporating KL divergence was employed after imputing the missing data with deep generative adversarial networks. This framework achieved a high prediction accuracy. Finally, Shankar and Arasu (2023) explored hidden temporal and periodic patterns in AQI data, demonstrating that hybrid models, such as CNN-LSTM and GRU, integrated with Empirical Mode Decomposition, surpassed their standalone counterparts in predictive performance. Collectively, these studies underscore the effectiveness of hybrid architectures in enhancing the accuracy and robustness of AQI prediction by leveraging a blend of preprocessing techniques and sophisticated deep learning models.

A Transformer model was proposed for $PM_{2.5}$ prediction at 12 Beijing sites and performed better than the CNN-LSTM-Attention model, with EVS, MAE, MSE, and R^2 being improved by 12%, 9%, 6%, and 30%, respectively (Cui et al. 2023). It successfully captured short-term meteorological changes and long-term seasonal trends and has an excellent ability to model long-range dependencies for air quality forecasting. This research (Wu et al. 2024) presented a partial least squares VAER (PLS-VAER) model for indoor air quality prediction, in which PLS extracts latent variables to enhance the VAER input and improve accuracy. The model outperformed conventional approaches and provided a robust and eco-friendly solution for indoor air quality monitoring and parameter optimization. In contrast to the current Bi-LSTM and hybrid models that utilize uniform sets of features across regions, this study focuses on station-specific feature engineering, adapting inputs for each site's environment and geography. This enables the capture of localized drivers of pollution, including coastal influence and landfill distance, which improves predictive performance and interpretability over traditional approaches.

3. TEMPORAL AND CORRELATION ANALYSIS

This section discusses the temporal correlation of AQI values, followed by seasonal and diurnal oscillations between the AQI and factors driving AQI levels, such as air contaminants and meteorological conditions. This research draws on data from two sources: 1. CPCB site, 2. OpenWeatherMap API. Atmospheric air pollutant data for the two stations, Colaba and Deonar, were obtained from the CPCB portal (Central Pollution Control Board 2024) for the periods of July 2019 to March 2024 for Colaba station and November 2020 to March 2024 for Deonar station. Meteorological data for the same stations and timestamps were captured via the OpenWeatherMap API (OpenWeather 2024).

3.1 Temporal Analysis

There was a strong autocorrelation between the AQI values, as shown in Fig. 1. Present AQI values are associated with past AQI values; that is, the present values are dependent on past AQI values. In Fig. 1, for both stations, all points are centered around the diagonal, which indicates that the AQI time series has the same descriptive properties, such as mean, variance, and standard deviation, over different time instances, which shows that both time series are stationary in nature (Douglas et al. 2015).

3.2 Correlation Analysis

All six atmospheric pollutants were positively associated with AQI levels in the air, as shown in Fig. 2. Particulate matter, PM2.5, and PM10 had nearly a perfect positive correlation, whereas NOx and SO_2 also exhibited a strong positive association with AQI levels. For ozone, there was a location-specific dependency; that is, it showed a modest positive influence on AQI levels for the Deonar station but not much for the Colaba station.

Similarly, the Pearson Coefficient was used to understand the relationship between meteorological parameters and AQI levels. For both stations in the Figs. 3 and 4, it is apparent that temperature, dew point, minimum and maximum temperature in a day, humidity, wind speed, and wind direction have a negative association with AQI levels, whereas diurnal temperature and pressure have a positive impact on AQI levels. It was observed that out of all meteorological parameters, pressure, humidity, and dew point had a strong association with the target variable.

3.3 Statistical Analysis

Table 1 presents the summary statistics of AQI values at Colaba and Deonar stations, showing dispersion in terms

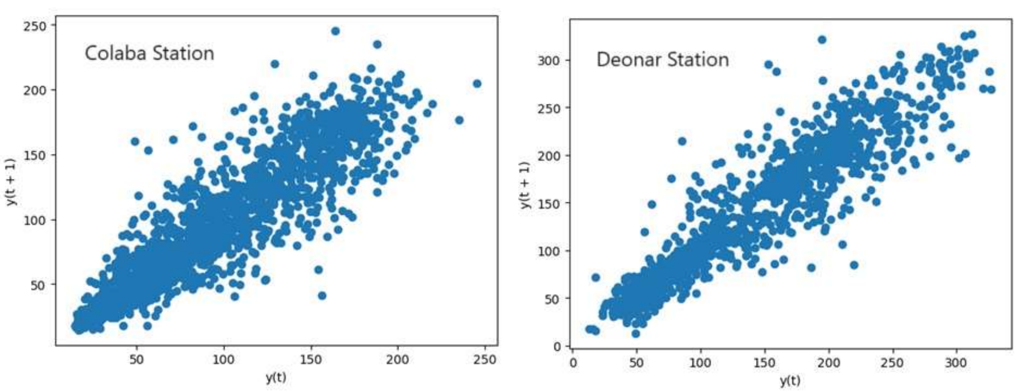

Fig. 1: Auto-lag plots of AQI values for the Colaba and Deonar Stations.

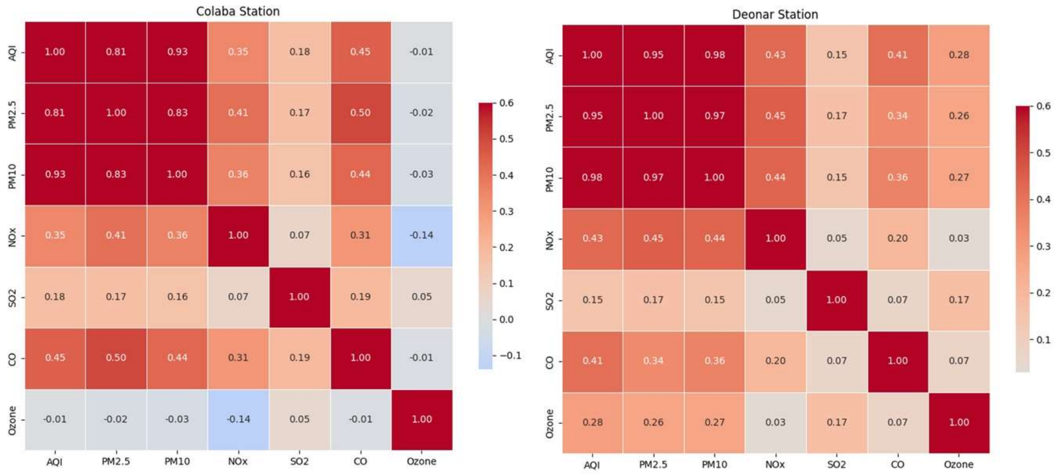

Fig. 2: Heatmap of Pearson correlation between atmospheric pollutants and AQI values for Colaba and Deonar Stations.

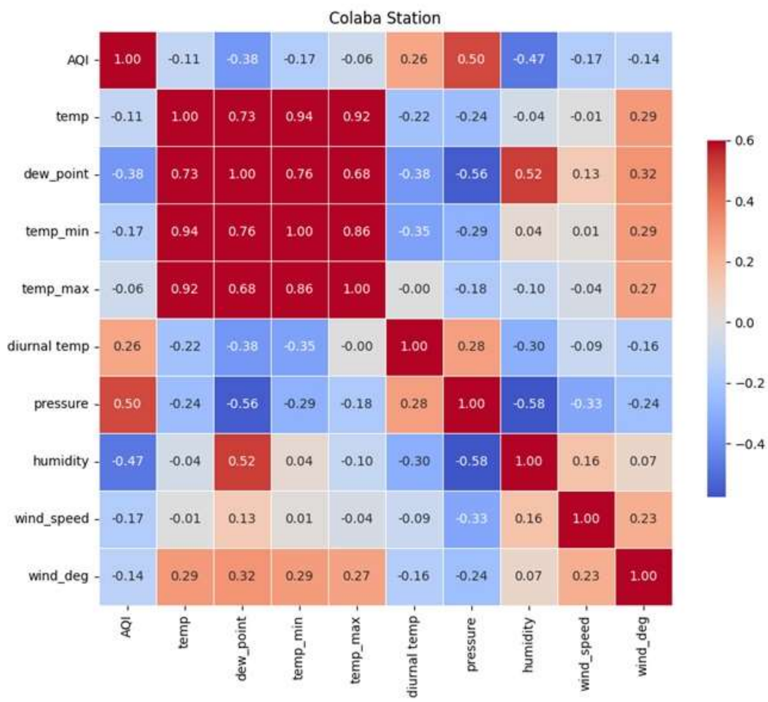

Fig. 3: Heatmap of Pearson correlation between meteorological factors and AQI values for the Colaba Station.

of variance and standard deviations, seasonal variation by calculating winter and summer means, and diurnal variation with day and night means for winter and summer seasons. Both time series are volatile in nature, with Deonar showing a higher dispersion or spread.

In all respects, Deonar always showed higher AQI values in both seasons and at all times of the day, indicating a generally poorer air quality compared to Colaba. In winter, both stations had higher readings at night than during the day. During summer, both stations indicated higher values during the day than at night, with Colaba showing more pronounced day–night differences, indicating that the time of day could have a notable impact on air quality. To understand the variability in AQI levels, both seasonal and diurnal variations must be captured precisely.

4. MATERIALS AND METHODS

Fig. 5 shows a detailed diagram of the AQI prediction workflow using deep learning models. There are four stages: Data Collection, Data Pre-processing, Feature Engineering,

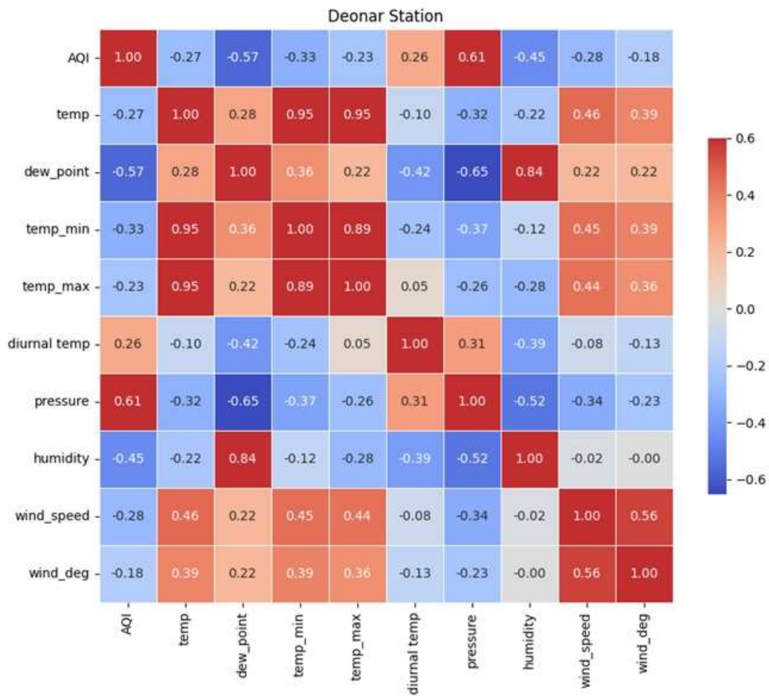

Fig. 4: Heatmap of Pearson correlation between meteorological factors and AQI values for the Deonar Station.

Table 1: Statistical summary of AQI values for Colaba and Deonar Stations.

Stations	Dispersion		Seasonal Variation		Diurnal Variations			
	Variation	Standard Deviation	Winter Mean	Summer Mean	Winter Day Time Mean	Winter Night Time Mean	Summer Day Time Mean	Summer Night Time Mean
Colaba	4532.95	67.32	73.06	114.16	70.45	76.03	144.26	116.32
Deonar	7382.04	85.92	202.32	120.18	199.65	205.81	190.62	120.51

and Model Development. In stage 1, data were collected from multiple sources, including the concentrations of various atmospheric pollutants, such as particulate matter with 15 and gaseous pollutants, along with meteorological parameters, namely, temperature, wind speed, wind direction, atmospheric pressure, dew point, and humidity.

A combined dataset feature was created using pollutant levels and meteorological factors. The data were for hourly records, with Colaba having 41,617 data points from July 01, 2019, to March 31, 2024, and Deonar having 29,929 data points from November 01, 2020, to March 31, 2024, with the gap being due to non-availability of data on the CPCB portal for some dates. The AQI calculator (Central Control Room for Air Quality Management Delhi NCR 2024) was used to calculate the AQI value, which was also appended to the dataset. Data pre-processing is stage 2, which includes multiple steps, such as filling missing values using a time-based mean interpolator, followed by detecting irregularities in the data using interquartile ranges and filling those irregularities with a time-based mean imputer. The TBMI algorithm imputes step by step, first computing the group means by time, month, day, and year, and then imputing the missing values with the calculated group means. The two-step anomaly treatment used in this study began with the inter-quantile range (IQR) method at the 15th and 85th percentiles to mark values beyond this range. This method identifies medium-sized anomalies without a high degree of sensitivity to extreme values. Before feeding for training, the data were standardized and normalized, which is equivalent to adjusting the values to a common scale, ensuring uniformity and comparability across the dataset.

The next stage is feature engineering, where multiple feature sets are created based on the results of Pearson's correlation matrix. These experiments and combinations were created to understand how different deep learning topologies interact with different feature vectors. For this study, three different feature vectors, namely, Best Pollutant parameters, Best Meteorological parameters, and

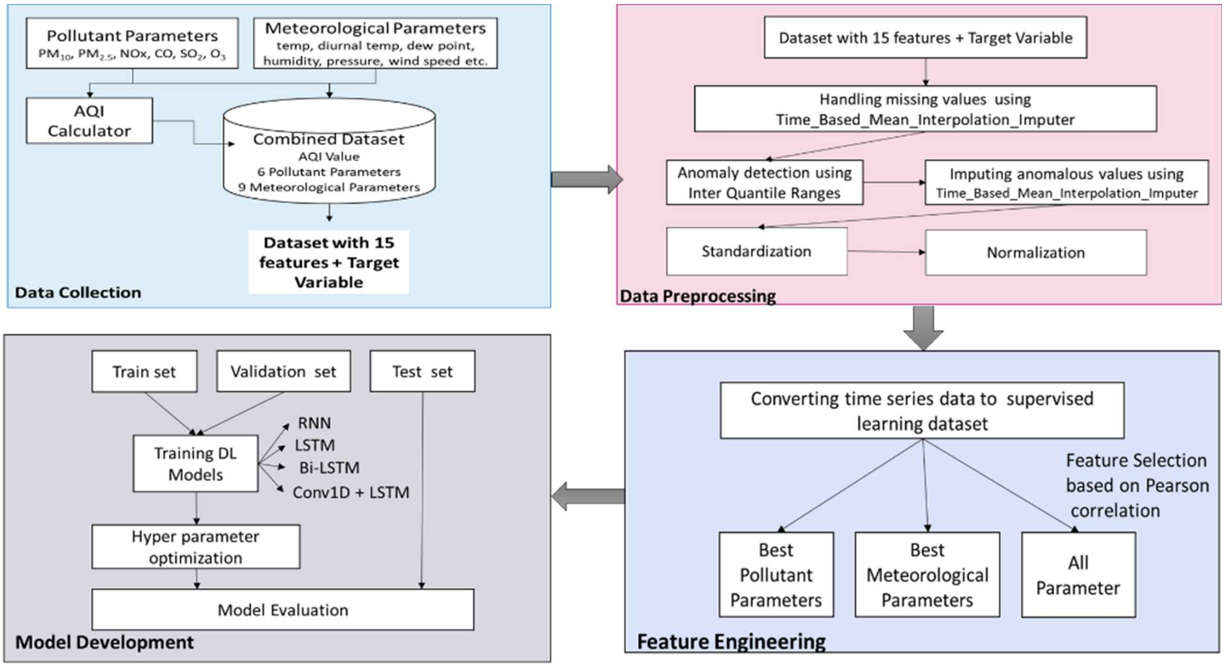

Fig. 5: Comprehensive deep learning workflow for accurate AQI prediction.

Table 2: Selected Features for AQI Prediction at Colaba and Deonar.

Stations	Colaba	Deonar
Best Pollutant Parameters	PM_{10}, $PM_{2.5}$, NOx, CO	PM_{10}, $PM_{2.5}$, NOx, CO, Ozone
Best Meteorological Parameters	Pressure, Humidity, Dew Point, Diurnal temperature	Pressure, Humidity, Dew Point, Temperature, Wind Speed
All Parameters	6 Pollutant Levels + 9 Meteorological Parameters + AQI (Target Variable)	6 Pollutant Levels + 9 Meteorological Parameters + AQI (Target Variable)

all 15 parameters, were used, as shown in Table 2. The last stage of this research included the development of a model, which involved splitting the data into multiple sets, namely, training, validation, and test sets, with a 70:20:10 ratio for both stations. Using optimized hyperparameter settings along with the training and validation sets, multiple deep learning architectures, namely, RNN, LSTM, Bidirectional LSTM, and Conv1D+LSTM, as shown in Fig. 6, were trained using MSLE as the loss function. For all these architectures, the past week of data was used to make a single-step ahead AQI prediction. The shape of the input was different for different architectures, depending on the feature set and location.

The first architecture, that is, the RNN architecture, has an input layer followed by a couple of RNN layers and eventually a dense layer before the final output layer. The RNN layer processes sequential data and captures temporal dependencies using hidden states over multiple time steps. The dense layer was used to project the outputs of the RNN into the desired output dimension. RNNs suffer from the problem of diminishing gradients (Khan et al. 2023). The second architecture, which is a stacked LSTM, is an enhancement over the simple RNN, where RNN cells are replaced by LSTM cells. LSTM cells can reduce the vanishing gradient issue with the help of a gating mechanism, where different gates decide which information to pass and which information to remove (Le et al. 2020).

To avoid overfitting, dropout regularization is used, and stacked configurations can help learn more complicated patterns by increasing their depth. The third architecture studied is a stacked Bidirectional LSTM, where the input sequences are operated in both forward and backward directions. These models can perceive latent relationships between the input and target variables in both past and future contexts (Mandal & Sarode 2024a). The architecture is similar to the LSTM architecture, with one change: LSTM cells in the hidden layers are replaced with Bi-LSTM cells. Finally, a hybrid architecture uses convolution layers to extract local patterns and then feeds those local extracted

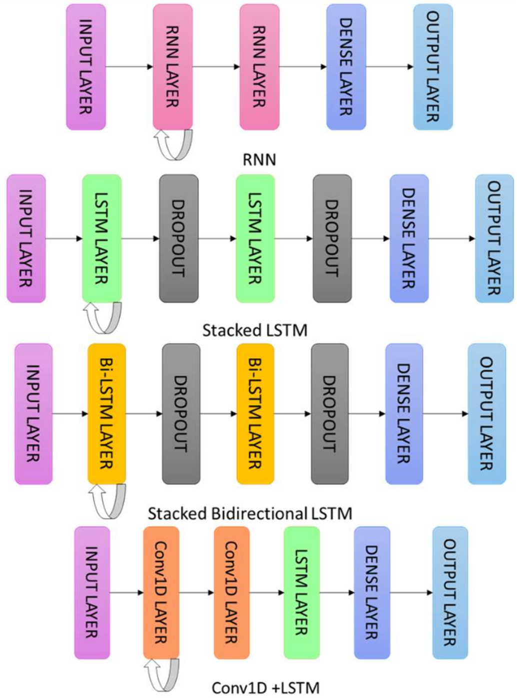

Fig. 6: Architectural diversity of deep learning models for AQI prediction.

patterns to LSTM layers to capture temporal relationships (Jiao et al. 2020). This hybrid model can take advantage of both learning topologies.

4.1 Evaluation Metrics

Multiple performance metrics were used to understand different aspects of the model's performance. These metrics provide a comprehensive evaluation of prediction models covering various facets, such as the magnitude of error in prediction and the extent of variation in the data captured.

RMSE: The Root Mean Squared Error (RMSE) is a measure of the average magnitude of the errors between the actual and predicted AQI values. It calculates the square root of the average of the squares of the differences between every actual value and its predicted value (Mandal & Sarode 2024b, Botchkarev 2019). Squaring gives more importance to larger errors than to smaller ones; hence, this metric is sensitive to the outliers. A lower RMSE value indicates that the average of the predicted values is closer to the actual value.

R^2 Score: The R^2 Score is the coefficient of determination (Plevris et al. 2022, Botchkarev 2019), which indicates how well the model fits the variability of the target variable. It provides a number showing the proportion of variance in the dependent variable that might be predicted from the independent variables. The R^2 score can take any value from 0, which means that the model does not explain any variability in the target variable, to 1, which would mean that it explains all the variability. A higher R^2 score indicates how well the model fits the data and, therefore, how well it predicts the AQI values based on the input features.

MAPE: MAPE stands for Mean Absolute Percentage Error, which is an accuracy metric calculated by taking the average absolute percentage difference between the predicted and actual values (Plevris et al. 2022, Botchkarev 2019). This will be useful when the accuracy of a model on different datasets or scales is compared. MAPE normalizes these errors. The lower the value of MAPE, the more accurate the predictions; the higher the value, the greater the error.

Another explanatory metric is EVS, which stands for Explained Variance Score (EVS) (Plevris et al. 2022), which indicates the proportion of variance in a target variable explained by the predicted values from a model. This lies between 0 and 1, where higher values indicate that more of the variability is accounted for by the model. An EVS of 1 would mean a perfect explanation of the variance by the model, and one close to 0 would mean very little explanatory power. The EVS indicates how well the model captures the underlying patterns in the data.

5. RESULTS AND DISCUSSION

5.1 Experimental Setup

For this experiment, every combination of deep learning networks along with different feature vectors was run for

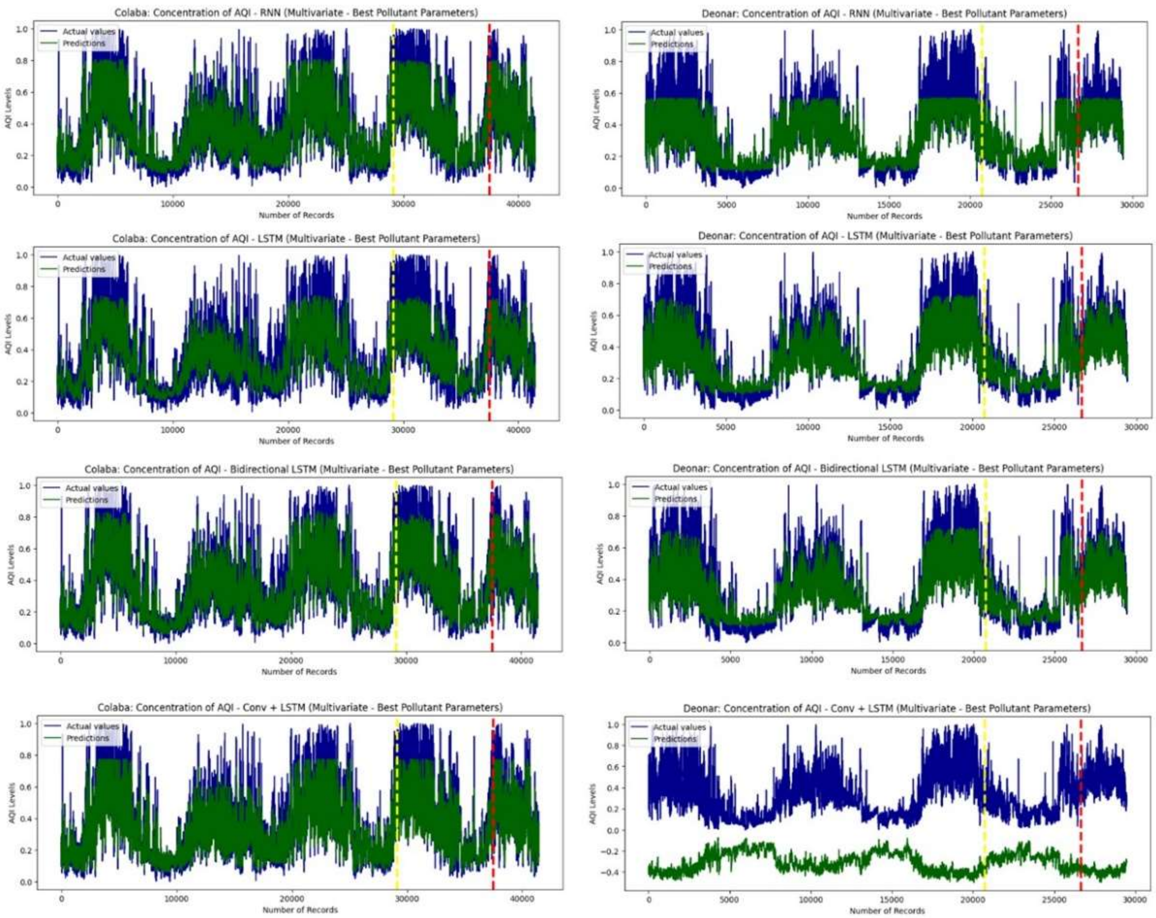

Fig. 7: Actual and predicted AQI values using various deep learning models with the Best Pollutant Feature Set.

50 epochs with a batch size of 256 samples. For recurrent hidden layers, the activation function used was ReLU, and the number of hidden units was 64, followed by 32. For the dense layer, linear activation was used with 16 and 1 neurons. For efficient training, the Mean Squared Logarithmic Error (MSLE) was used as the loss function, and the Adam optimizer with 0.001 as the learning rate was used for stable convergence with dropout regularization of 0.2.

5.2 Best Pollutant Feature Set

Fig. 7 shows the AQI prediction for both stations using multiple deep learning topologies for the best pollutant parameters. The actual values are shown in blue, and the predicted values are shown in green. 70% of the dataset was used for training, which is presented to the left of the yellow dotted line. Next, 20% of the data were used as the validation set, which is between the yellow and red dotted lines, and the remaining 10% were used as the test set. It was observed that with pollutant parameters as feature sets, better predictions were made for Colaba station than Deonar station, which might be due to the high spread of AQI values for Deonar station. For the Colaba station, the RNN and Bidirectional LSTM made slightly better AQI predictions, whereas for the Deonar station, the LSTM provided significantly better predictions, and it is noticeable that Conv1D + LSTM with pollutant feature set failed.

5.3 Best Meteorological Feature Set

Fig. 8, which demonstrates the actual and predicted AQI values for all deep learning architectures for both stations using meteorological features, shows that every model captured the variation in AQI values more precisely with this set, indicating better prediction capability. By closely examining the plots and using climatic parameters, the Conv1D + LSTM model resulted in a marginal improvement in making predictions for the Colaba station, whereas for the Deonar station, LSTM again outperformed.

5.4 Combined Feature Set with All Parameters

For the combined feature set with a total of 15 features and

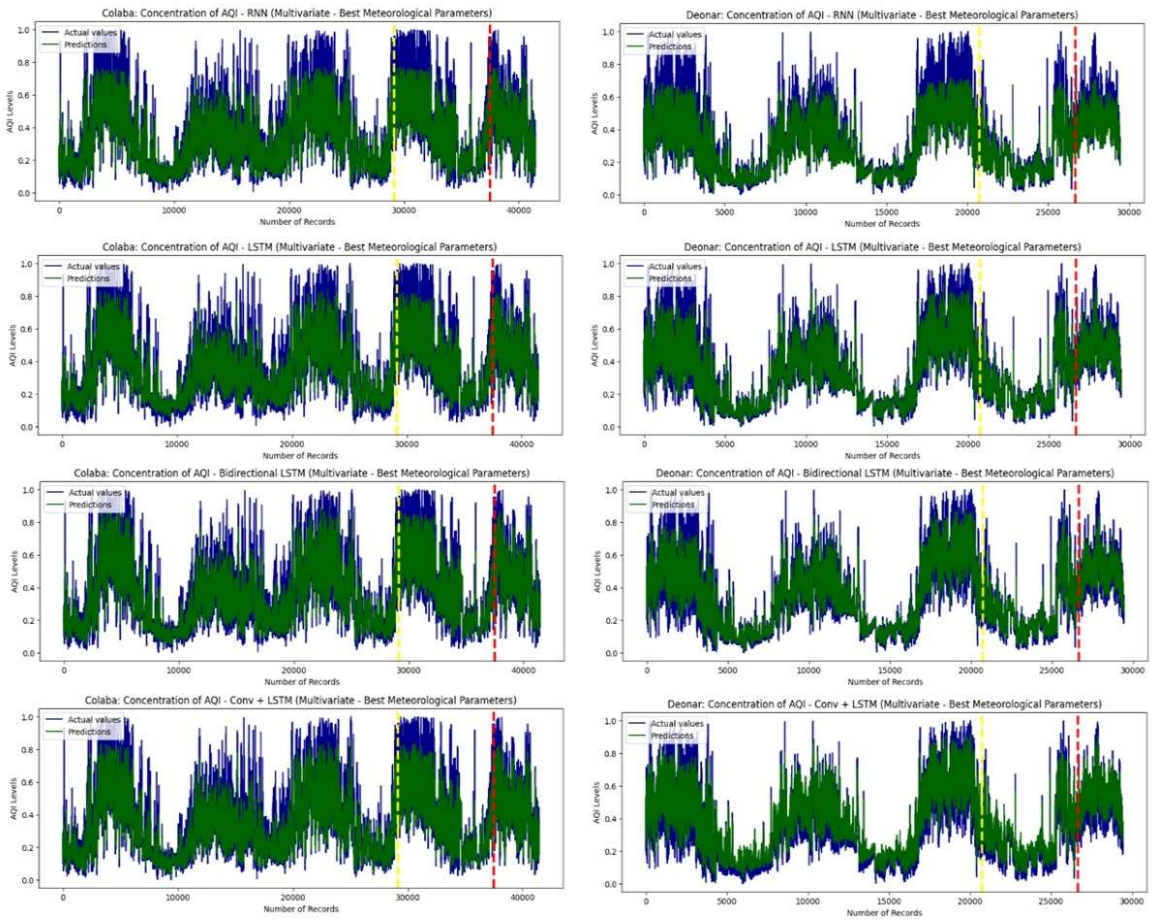

Fig. 8: Actual and predicted AQI values using various deep learning models with the Best Meteorological Feature Set.

target variables, the actual and predicted values for all deep learning architectures used for both stations are shown in Fig. 9. For the Colaba station, similar to the meteorological feature set, the Conv1D + LSTM model made the best predictions with little improvement over the other models. Owing to the high dispersion in the data for the Deonar station, there were significant fluctuations in the outcomes of various models. It is evident that the RNN model has completely failed, and other models have resulted in similar results, with Bi-LSTM being slightly better.

5.5 Juxtaposed Analysis

The performance of various deep learning architectures for different input feature sets is tabulated in Tables 3 and 4 for the Colaba and Deonar stations, respectively. These tables include varying performance measures, namely, RMSE, R^2 Score, MAPE, and EVS.

Across Models: The Bi-LSTM and LSTM models exhibited consistent performance across different stations and feature sets. For Colaba, it provided the lowest RMSE and MAPE with several feature sets and the highest R^2 Score and EVS, thus showing better prediction accuracy. At Deonar, the Bi-LSTM model performed well, especially with meteorological parameters, producing competitive results. Although Conv1D + LSTM has some bright spots, considering the overall model reliability and efficiency in minimizing errors and explaining enhanced variance, Bi-LSTM and LSTM are the best models, irrespective of the station or feature set.

Across feature sets: The meteorological feature set outperformed the others across both Colaba and Deonar stations. Indeed, models with meteorological parameters recorded the lowest RMSE and MAPE and higher values for the R^2 Score and EVS, thus showing better variance explanation. The discovery that meteorological factors affect the AQI more significantly than pollutant concentrations is important because these factors directly regulate pollutant spreading and settling. Increased temperature and humidity reduce the AQI by increasing atmospheric mixing and facilitating pollutant removal. Conversely, extended high-

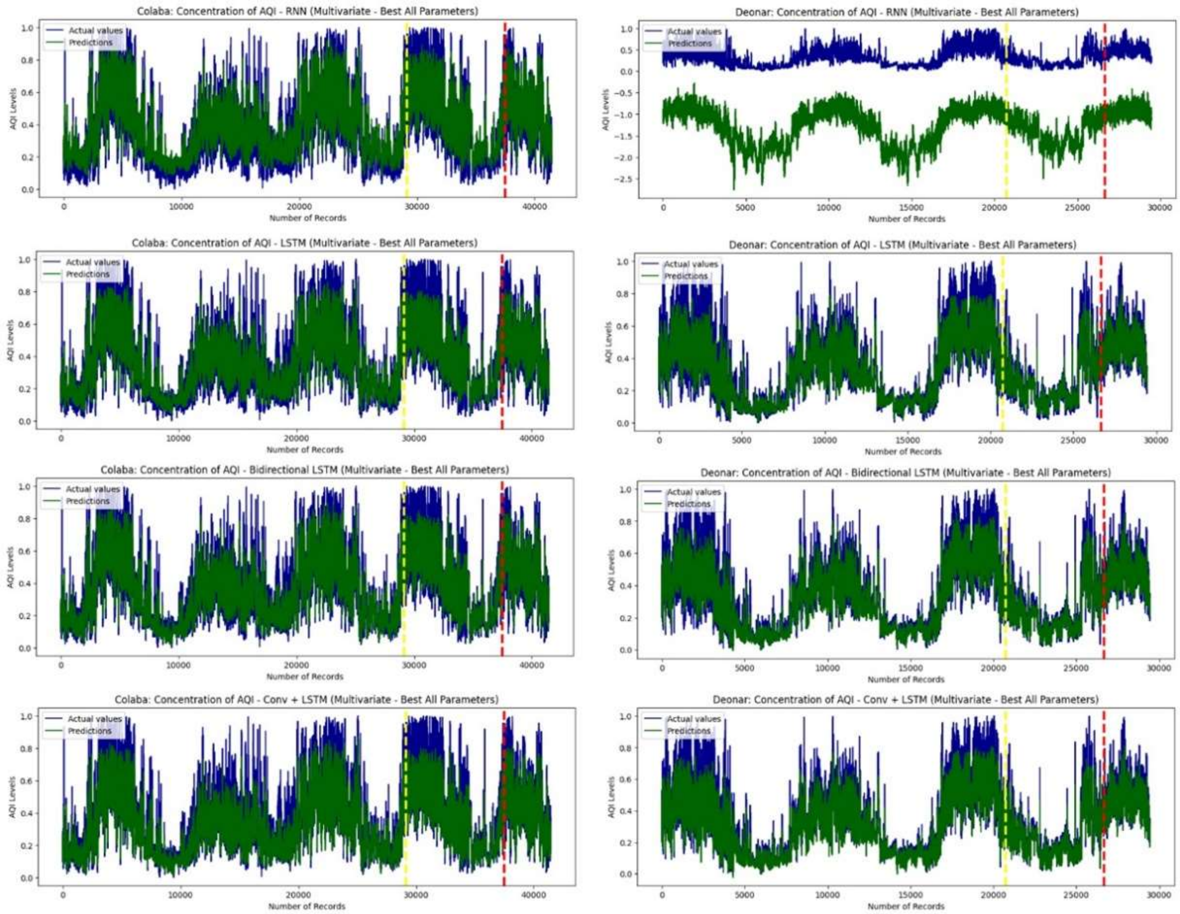

Fig. 9: Actual and predicted AQI values using various deep learning models with a Combined Feature Set with All Parameters.

Table 3: Performance Metrics of Various Models across Different Feature Sets for AQI Prediction for Colaba Station.

Feature Set	Models	RMSE	R^2 Score	MAPE	EVS
Pollutant Parameters	RNN	21.3230	0.7619	0.1268	0.8062
	LSTM	23.6293	0.6164	0.1502	0.7569
	Bi-LSTM	21.0952	0.7551	0.1260	0.8093
	Conv1D + LSTM	22.1642	0.6927	0.1334	0.8029
Meteorological Parameters	RNN	22.1090	0.6553	0.1312	0.7984
	LSTM	21.6329	0.6852	0.1322	0.7962
	Bi-LSTM	21.5187	0.7112	0.1327	0.7984
	Conv1D + LSTM	21.0152	0.7320	0.1238	0.8089
All Parameters	RNN	23.5945	0.6796	0.1545	0.8175
	LSTM	22.0414	0.7021	0.1400	0.8069
	Bi-LSTM	21.4796	0.7407	0.1329	0.8139
	Conv1D + LSTM	20.8821	0.7502	0.1246	0.8189

pressure conditions can elevate the AQI by generating stable atmospheric flows that retain pollutants in the lower layers of the atmosphere. Increased wind speeds lower the AQI by spreading out pollutants, whereas wind direction has a weaker, less predictable influence on air quality.

Table 4: Performance Metrics of Various Models across Different Feature Sets for AQI Prediction for Deonar Station.

Feature Set	Models	RMSE	R^2 Score	MAPE	EVS
Pollutant Parameters	RNN	30.6543	0.1596	0.1057	0.6140
	LSTM	33.7968	0.2082	0.1407	0.6980
	Bi-LSTM	34.9731	0.1983	0.1548	0.6725
	Conv1D + LSTM	362.7671	-689.501	2.3102	-0.347
Meteorological Parameters	RNN	30.1466	0.2836	0.1156	0.7034
	LSTM	26.2931	0.5135	0.0955	0.7116
	Bi-LSTM	27.8235	0.4953	0.1015	0.6892
	Conv1D + LSTM	34.4509	0.2719	0.1288	0.7180
All Parameters	RNN	588.7852	-96.7797	1.5506	-1.163
	LSTM	26.3419	0.5340	0.0956	0.7088
	Bi-LSTM	26.6565	0.5510	0.0977	0.7124
	Conv1D + LSTM	27.4466	0.5264	0.1028	0.7171

Across Locations: Models fare poorly for Deonar primarily because of the high variability in its AQI values, which makes prediction difficult. Such variability is presumably due to numerous factors specific to the area, including its location near mixed sources of pollution, such as landfills, industrial areas, and busy traffic routes. These sources create non-linear fluctuations in pollutant concentrations that are more difficult to incorporate using conventional modeling methods. Inconsistencies and gaps within the data also contribute to decreased model reliability. Aggregated, these factors produced more intricate and less predictable air quality patterns in Deonar than in Colaba, resulting in increased error measures and lower explained variance across models.

6. CONCLUSIONS

Air pollution is one of the prime factors of mortality across the globe and needs to be monitored and predicted accurately to reduce its impacts. Atmospheric pollutants, such as $PM_{2.5}$, PM_{10}, ammonia, SO_2, CO, NOx, and ozone, greatly affect human health. Climatic factors, including temperature, humidity, wind speed, and pressure, influence the AQI. The dataset from two stations, Colaba and Deonar, containing atmospheric pollutants and meteorological factors, was the research object of this study. Correlation studies show that AQI levels are strongly influenced by their past values and have a strong positive correlation with atmospheric pollutants and various meteorological conditions, such as temperature, pressure, and wind speed. Other climatic factors, such as humidity and dew point, were negatively associated with AQI levels.

For this study, based on the results of Pearson's correlation, multiple input feature sets, including the best pollutant set, best meteorological set, and combined feature sets, were created. For these sets, multiple deep learning topologies, such as RNN, LSTM, Bi-LSTM, and Conv1D + LSTM models, were implemented with optimized hyperparameters for both locations. The assessment metrics used were RMSE, R^2 Score, MAPE and EVS. For the Colaba station, RNN and Bi-LSTM outperformed the pollutant feature set, whereas Conv1D + LSTM showed improved results for the other two feature sets. For the Deonar station, LSTM produced the best results in all scenarios, and RNN with the combined feature set failed. Meteorological factors have enhanced the accuracy, as deep learning models are capable of learning more complex patterns that model nonlinear associations and interactions.

As different results were obtained for different locations, this ensures the need for more tailored location-specific models for more accurate predictions. This research advances by comparing multiple feature sets and multiple architectures with location-specific information. This information can be helpful for policymakers and environmental regulatory authorities to forecast episodes of poor air quality and release necessary early alerts and warnings, which are needed for public safety, support urban planning, and eventually reduce the burden on the healthcare infrastructure of the city.

Every region has sources of pollution, geographical features, climatic conditions, and human activities that affect air quality differently. Although generic models are helpful, they may lack such local variations, leading to less accurate predictions. In the future, studies can include other factors such as proximity to industrial areas, traffic density, vegetation cover, distance from landfill, and distance from the shoreline, which can all impact the dispersion and concentration of pollutants in the air. Attention-based systems, variational autoencoders, probabilistic graph networks with a greater number of training samples, and a more inclusive feature set can further enhance prediction accuracy.

7. ACKNOWLEDGMENTS

The authors recognize the National Environmental Engineering Research Institute (NEERI) in Nagpur, the Central Pollution Control Board, State Pollution Control Boards, and Pollution Control Committees for making data freely available for research purposes.

8. REFERENCES

Ayus, I., Natarajan, N. and Gupta, D., 2023. Comparison of machine learning and deep learning techniques for the prediction of air pollution: a case study from China. *Asian Journal of Atmospheric Environment*, 17(4), pp.1-8. [DOI]

Barve, A., Singh, V., Shrirao, S. and Bedekar, M., 2020. Air quality index forecasting using parallel dense neural network and LSTM cell. In: *2020 International Conference on INCET*, pp.1-4. [DOI]

Binbusayyis, A., Khan, M., Ahmed, A. and M.M., 2024. A deep learning approach for prediction of air quality index in smart city. *Discover Sustainability*, 5, p.89. [DOI]

Bizualem, B., Angassa, K., Shibeshi, N.T. and Feyisa, G.L., 2023. Assessing the status of air pollution related to traffic using dispersion modeling: the case of Megenagna, Addis Ababa, Ethiopia. *Bulletin of Atmospheric Science and Technology*, 4(1), p.9. [DOI]

Botchkarev, A., 2019. A new typology design of performance metrics to measure errors in machine learning regression algorithms. *Interdisciplinary Journal of Information, Knowledge, and Management*, 14, pp.45-76. [DOI]

Central Control Room for Air Quality Management Delhi NCR, 2024. Air Quality Index Information - AQI Calculator. Retrieved June 29, 2024, from https://app.cpcbccr.com/ccr_docs/AQI%20-Calculator.xls

Central Pollution Control Board (CPCB), 2024. CPCB Home Page. Retrieved June 24, 2024, from https://app.cpcbccr.com/AQI%20 India/

Cui, B., Liu, M., Li, S., Jin, Z., Zeng, Y. and Lin, X., 2023. Deep learning methods for atmospheric PM2.5 prediction: a comparative study of transformer and CNN-LSTM-attention. *Atmospheric Pollution Research*, 14(9), p.101833. [DOI]

Cui, H., Ma, R. and Gao, F., 2018. Relationship between meteorological factors and diffusion of atmospheric pollutants. *Chemical Engineering Transactions*, 71, pp.1417-1422. [DOI]

DeVries, R., Kriebel, D. and Sama, S., 2017. Outdoor air pollution and COPD-related emergency department visits, hospital admissions, and mortality: a meta-analysis. *COPD*, 14(1), pp.113-121. [DOI]

Douglas, C., Cheryl, L. and Murat, K., 2015. *Introduction to Time Series Analysis and Forecasting* (2nd ed.). Wiley, pp.450.

Fang, C., Lu, H., Li, G., Sun, D. and Miao, Z., 2015. Estimating the impact of urbanization on air quality in China using spatial regression models. *Sustainability*, 7(11), pp.15570-15592. [DOI]

Gou, A., Tan, G. and Ding, X., 2023. Urban-rural difference in the lagged effects of PM2.5 and PM10 on COPD mortality in Chongqing, China. *BMC Public Health*, 23, p.1270. [DOI]

Gupta, S., Mohta, Y., Heda, K., Armaan, R., Valarmathi, B. and Ganeshan, A., 2023. Prediction of air quality index using machine learning techniques: a comparative analysis. *Journal of Environmental and Public Health*, pp.1-26. [DOI]

Jian, Z., Cai, J. and Chen, R., 2024. A bibliometric analysis of research on the health impacts of ozone air pollution. *Environmental Science and Pollution Research*, 31, pp.16177-16187. [DOI]

Jiao, L. and Zheng, W., 2020. Multi-step traffic flow prediction method based on Conv1D + LSTM. In: *2020 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp.113-118. [DOI]

Kara, Y., Şevik, S.Y. and Toros, H., 2024. Comprehensive analysis of air pollution and the influence of meteorological factors: a case study of Adiyaman province. *Environmental Monitoring and Assessment*, 196, pp.1-12. [DOI]

Karthick, K., Aruna, S.K., Dharmaprakash, R. and Ravindiran, G., 2024. Integrating machine learning techniques for Air Quality Index forecasting and insights from pollutant-meteorological dynamics in sustainable urban environments. *Earth Science Informatics*, *17*(4), pp.3733-3748. [DOI]

Khan, D., Patankar, A., Motwani, R. and Ailsinghani, A., 2023. Investigating efficacy of RNN and its variants for AQI prediction. In: *2023 International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT)*, pp.1303-1307.

Kleingchuay, W., Srimanus, W., Srimanus, R. et al., 2022. The association of meteorological parameters and AirQ+ health risk assessment of PM2.5 in Ratchaburi province, Thailand. *Scientific Reports*, 12, p.12971. [DOI]

Le, V.-D., Bui, T.-C. and Cha, S.-K., 2020. Spatiotemporal deep learning model for citywide air pollution interpolation and prediction. In: *2020 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp.55-62. [DOI]

Liu, Y., Wang, P. and Li, Y., 2022. Air quality prediction models based on meteorological factors and real-time data of industrial waste gas. *Scientific Reports*, 12, p.9253. [DOI]

Lu, Y. and Li, K., 2023. Multistation collaborative prediction of air pollutants based on the CNN-BiLSTM model. *Environmental Science and Pollution Research*, 30, pp.92417-92435. [DOI]

Mandal, A. and Sen, R., 2024. Medium-term AQI prediction in selected areas of Bangladesh based on bidirectional GRU network model. *SN Computer Science*, 5, p.627. [DOI]

Mandal, N. and Sarode, T., 2024a. A framework for cloud cover prediction using machine learning with data imputation. *International Journal of Electrical and Computer Engineering (IJECE)*, 14(1), pp.600-607. [DOI]

Mandal, N. and Sarode, T., 2024b. A framework for cloud cover prediction using machine learning with data imputation. *International Journal of Electrical and Computer Engineering (IJECE)*, 14(1), pp.600-607. [DOI]

Méndez, M., Merayo, M.G. and Núñez, M., 2023. Machine learning algorithms to forecast air quality: a survey. *Artificial Intelligence Review*, 56, pp.10031-10066. [DOI]

Naz, F., McCann, C., Fahim, M., Cao, T.-V., Hunter, R., Viet, N.T. and Duong, T.Q., 2023. Comparative analysis of deep learning and statistical models for air pollutants prediction in urban areas. *IEEE Access*, 11, pp.64015-64025. [DOI]

Nikpour, P., Shafiei, M. and Khatibi, V., 2024. GELATO: a new hybrid deep learning-based informer model for multivariate air pollution prediction. *Environmental Science and Pollution Research*, 31, pp.29870-29885. [DOI]

OpenWeather, 2024. OpenWeather Homepage. Retrieved April 21, 2024, from https://openweathermap.org

Petrus, M., Popa, C. and Bratu, A.-M., 2022. Ammonia concentration in ambient air in a periurban area using a laser photoacoustic spectroscopy detector. *Materials*, 15(9), p.3182. [DOI]

Plevris, V., Solorzano, G., Bakas, N. and Ben Seghier, M.E.A., 2022. Investigation of performance metrics in regression analysis and machine learning-based prediction models. In: *Proceedings of the 8th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS 2022)*, p.155. [DOI]

Poyyamoli, G. and Boss, C., 2014. Assessment of air pollution and its impacts near municipal solid waste dumping site Kammiyampet, Cuddalore, India. *International Journal of Innovative Research in Science, Engineering and Technology*, 3(5), pp.12588-12593.

Rangaswamy, P.S., Vasareddy, A., Saiteja, B., Sreeja, B. and Chandana, G., 2022. Impact assessment on air quality around integrated municipal solid waste management plant in Hyderabad. *International Journal of Advanced Research in Science, Communication and Technology*, 2(1), pp.666-677. [DOI]

Saadi, D., Tirosh, E. and Schnell, I., 2021. The relationship between city size and carbon monoxide (CO) concentration and their effect on heart rate variability (HRV). *International Journal of Environmental Research and Public Health*, 18(2), p.788. [DOI]

Sajjad Abdollahpour, S., Qi, M., Le, H.T.K. and Hankey, S., 2024. Urban spatial structure and air quality in the United States: evidence from a longitudinal approach. *Environment International*, 190, p.108871. [DOI]

Schürholz, D., Kubler, S. and Zaslavsky, A., 2020. Artificial intelligence-enabled context-aware air quality prediction for smart cities. *Journal of Cleaner Production*, 271, p.121941. [DOI]

Shankar, L. and Arasu, K., 2023. Deep learning techniques for air quality prediction: a focus on PM2.5 and periodicity. *Migration Letters*, 20(S13), pp.468-484. [DOI]

Shi, T., Li, P. and Yang, W., 2023. Application of TCN-BiGRU neural network in PM2.5 concentration prediction. *Environmental Science and Pollution Research*, 30, pp.119506-119517. [DOI]

Song, Q., Zou, J. and Xu, M., 2023. Air quality prediction for Chengdu based on long shortterm memory neural network with improved jellyfish search optimizer. *Environmental Science and Pollution Research*, 30, pp.64416-64442. [DOI]

Sterling, M., 2017. Air quality: population density and transportation emissions. *The Equilibrium*, 3(1), pp.1-8. [DOI]

Sun, R., Zhou, Y., Wu, J. and Gong, Z., 2019. Influencing factors of PM2.5 pollution: disaster points of meteorological factors. *International Journal of Environmental Research and Public Health*, 16(20), p.3891. [DOI]

Thakur, N., Karmakar, S. and Shrivastava, R., 2023. Hybrid deep learning algorithms for predicting air quality index using dimension reduction technique in search of precise results. *International Journal of Information Technology*, 15, pp.3181-3187. [DOI]

Tian, J., Fang, C., Qiu, J. and Wang, J., 2021. Analysis of ozone pollution characteristics and influencing factors in Northeast Economic Cooperation Region, China. *Atmosphere*, 12(7), p.843. [DOI]

Varaprasad, V., Kanawade, V.P. and Narayana, A.C., 2024. Association between sea-land breeze and particulate matter in five coastal urban locations in India. *Science of the Total Environment*, 913, p.169773. [DOI]

Wang, P., Bi, X. and Zhang, G., 2023. A new hybrid PM volatility prediction model based on EMD and machine learning algorithms. *Environmental Science and Pollution Research*, 30, pp.82878-82894. [DOI]

Wang, S., Ren, Y. and Xia, B., 2023. Estimation of urban AQI based on interpretable machine learning. *Environmental Science and Pollution Research*, 30, pp.1-13. [DOI]

Wu, H., Hong, S., Hu, M., Li, Y. and Yun, W., 2022. Assessment of the factors influencing sulfur dioxide emissions in Shandong, China. *Atmosphere*, 13(1), p.142. [DOI]

Wu, Q., Geng, Y., Wang, X. et al., 2024. A novel deep learning framework with variational auto-encoder for indoor air quality prediction. *Frontiers in Environmental Science & Engineering*, 18(8), pp.1-12. [DOI]

Yan, H., Li, Q. and Feng, K., 2023. The characteristics of PM emissions from construction sites during the earthwork and foundation stages: an empirical study evidence. *Environmental Science and Pollution Research*, 30, pp.62716-62732. [DOI]

Yan, M. and Li, T., 2024. A review of the interactive effects of climate and air pollution on human health in China. *Current Environmental Health Reports*, 11, pp.102-108. [DOI]

Zhan, C., Li, S., Li, J., Guo, Y., Wen, Q. and Wen, W., 2020. Prediction of air quality in major cities of China by deep learning. In: *Proceedings of the 2020 16th International Conference on Computational Intelligence and Security (CIS)*, pp.68-72. [DOI]

Zhang, C. and Zhang, L., 2024. The relationship between toxic air pollution, health expenditure, and economic growth in the European Union: fresh evidence from the PMG-ARDL model. *Environmental Science and Pollution Research*, 31, pp.21107-21123. [DOI]

Zhang, Y. and Jiang, W., 2018. Pollution characteristics and influencing factors of atmospheric particulate matter (PM2.5) in Chang-Zhu-Tan area. *IOP Conference Series: Earth and Environmental Science*, 108, p.042047. [DOI]