

RESEARCH

Open Access



Deep learning-based AQI forecasting: a CNN-LSTM model with visual insights from SHAP-LIME and PDP

Ekata Mohapatra¹ , Mira Das¹ and Smita Rath^{2*}

*Correspondence:

Smita Rath
smitarath@soa.ac.in

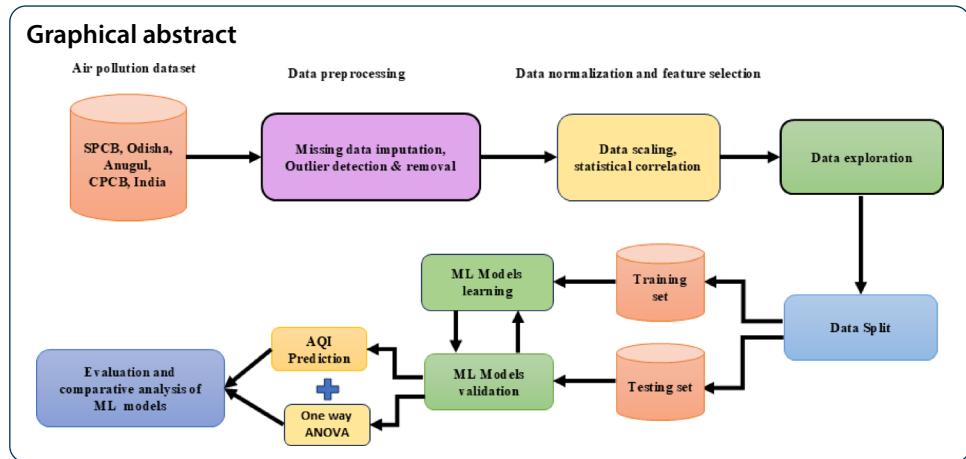
¹Environmental Science Programme, Department of Chemistry, ITER, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, Odisha, India

²Department of Computer Science and Information Technology, ITER, Siksha 'O' Anusandhan Deemed to be University, Bhubaneswar, Odisha, India

Abstract

Precise air quality forecasting is essential for assessing environmental health risks and facilitating timely intervention measures. This paper introduces a hybrid deep learning framework for multi-step Air Quality Index (AQI) forecasting based on multivariate time series data of Anugul, Odisha, and Cities of India. The proposed architecture incorporates convolutional neural networks (CNN) to identify spatial correlations among contaminants and long-short-term memory (LSTM) networks to identify long-term trends. The model exhibited robust predictive capability for Anugul city, Odisha, and cities of India with a Mean Squared Error (MSE) of 130.66, 1192, 130.66, Root Mean Squared Error (RMSE) of 11.40, 34.53, 37.75, and Mean Absolute Error (MAE) of 8.38, 27.65, 23.81, respectively, which is better than the accuracy of customary statistical and isolated deep learning models. To improve interpretability, Shapley Additive Explanations (SHAP), LIME (Local Interpretable Model-Agnostic Explanations), and Partial Dependence Plots (PDPs) were integrated, allowing for analysis of the contribution of each pollutant to the AQI predictions. The tools of interpretability aid in identifying environmental and industrial variables impacting air quality, enabling greater insights and giving more confidence in model outcomes. One-way ANOVA suggests that CNN-LSTM is a powerful model for AQI prediction when trained and validated separately for each data set. The results of p-values for all three metrics (MSE, RMSE, MAE) are below 0.5, which suggests the model's performance is highly region-dependent. The proposed hybrid CNN-LSTM method provides stable AQI prediction and enables interpretable decision-making via model explainability, providing application value for environmental monitoring and policymaking.

Keywords Air quality index (AQI), Long short-term memory (LSTM), Convolutional neural networks (CNN), Shapley additive explanations (SHAP), Local interpretable model-agnostic explanations (LIME), Partial dependence plots (PDP)



1 Introduction

Urban air pollution has become a significant issue and a hurdle to sustainable development and the construction of a green civilization [22]. Air quality directly affects the health and livelihood of a human being [24]. The Air Quality Index (AQI) is a vital indicator for determining air quality. Thus, accurately predicting future AQI changes from historical data is of great significance [60]. Artificial intelligence is a revolutionary new technology that makes improvements in deep learning algorithms, including Deep Belief Networks, Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) [21]. These algorithms have broad usage across pattern recognition, object detection, natural language, and image classification. As analytics demands increase in volume and sophistication, a solitary neural network model proves inadequate. Hybrid network models are superior methods [37]. Though widely used for time-series prediction, long Short-Term Memory (LSTM) networks encompass parameters that demand heavy computations and longer training durations. However, CNNs have the drawback of overfitting with small data. However, CNNs are capable of refining the feature extraction from input data. Comparative experiments show that CNN and LSTM models enhance the correctness of AQI estimation [10]. Air quality monitoring is crucial for public health and environmental management, especially in highly populated urban regions with pollution from traffic, industrial activities, and residential sources that degrade air quality [29]. Conventional methods have relied on fixed observation stations, which, even if well situated, often have limited spatial coverage and do not usually provide real-time data for timely interventions during pollution events [39]. Mishra and Gupta did a comparative analysis of statistical models and some other models for air quality forecasting [43].

Levels of air pollution show discernible seasonal patterns, with changes primarily driven by meteorological factors and human activities [17]. During winter, low temperatures, gentle winds, and atmospheric inversion retain pollutants close to the ground, leading to maximum pollution and poor air [62]. Post-monsoon and summer seasons show moderate air pollution with high temperatures, resulting in dilution of pollutants. However, increased dust levels and emissions from construction activities and agricultural practices contribute to air pollution [3]. The study indicates combining multiple models to analyse complex systems for forecasting AQI concentration levels based on available data from the Odisha Pollution Control Board (OPCB) and the Central Pollution Control Board (CPCB) of India. The model has been evaluated with Anugul

city pollutant data and assessed its performance relative to conventional statistical and machine learning approaches using standard cross-validation metrics. Air pollution is now a worldwide environmental problem because of its harmful effects on climate change, the environment, and human health [54]. Air Quality Index (AQI) is an indicator most frequently applied for describing the general air quality and estimating possible adverse health effects of pollutants like nitrogen dioxide (NO_2), ozone (O_3), sulphur dioxide (SO_2), and particulate matter ($\text{PM}_{2.5}$, PM_{10}) [48]. Accurate and timely predictions of AQI have various important implications, including mitigating the health effects of air pollution, supporting the formulation of effective environmental policy, and issuing timely public health warnings [45]. In recent years, high-level data-driven modelling techniques for forecasting and environmental study possess superior capabilities to represent the temporal patterns and intricate nonlinear correlations [12]. Specifically, Long Short-Term Memory (LSTM) networks dominate long-range dependencies on sequential data, and Convolutional Neural Networks (CNNs) can successfully execute the local spatial vs. temporal characteristics on high-dimensional data [56]. CNN-LSTM hybrid models take advantage of these features and can be very successful for AQI prediction problems, which require short-term variation and long-term trends [4, 5]. exhibited that their hybrid CNN-LSTM model provides better forecasting accuracy than all the traditional models [50]. Demonstrate that the hybrid CNN-LSTM model performs better than other machine learning methods like transformers, GRU, Graph Neural Networks, TabNet, and similar approaches in the case of AQI forecasting. It accurately identifies spatial and temporal patterns within air quality data.

Even though deep learning models are very accurate, they are opaque, making it difficult to understand and rely on when making decisions. Partial Dependence Plots (PDPs) are receiving much attention to avoid this. PDPs assist in visualizing the marginal impact of one feature on model predictions and give insights into how various pollutants or weather conditions impact the predicted AQI [9]. Including PDP visualization in AQI prediction models makes the models transparent and allows for data-driven environmental policy interventions [49]. This research discusses creating a hybrid CNN-LSTM model for AQI prediction and its performance evaluation and explainability based on PDP visualizations. We compute SHAP (Shapley Additive Explanations) values for each method to depict feature contribution and relevance to model prediction, offering a unified framework of interpretability. LIME offers local explanations. This approach makes comprehending estimates for specific scenarios easier and provides insights into local decision boundaries [26]. The objective is to obtain predictive accuracy with explainability, linking intricate model behaviour to helpful environmental understanding.

2 Literature review

Air pollution impacts climate, ecosystems, virus spreading, human health, and sustainable socio-economic development. The alteration of air pollution time series is nonlinear and influenced by meteorological conditions [20]. Removal and dispersion of air pollutants are governed by the atmospheric conditions for diffusion and precipitation [19].

The air quality index (AQI) is a comprehensive index free of units to represent air quality quantitatively. It is evaluated using the new ambient air quality criteria and can incorporate all pollutants (PM_{10} , $\text{PM}_{2.5}$, O_3 , CO , NO_2 , SO_2). Significantly, the AQI is a crucial measure of the intimate relationship between air quality and human health, and

it has become a widely accepted health indicator. The higher the score, the greater the health hazards and the necessity of preventative action [61]. The AQI aims to promote public education about the harmful health effects of local air pollution. In Indian cities, air pollution has dramatically increased [25].

Researchers commonly divide AQI forecasting methods into (i) ML-based methods, (ii) DL-based methods, and (iii) hybrid methods.

Computer software, which is known as an ML-based technique, can recognize patterns in historical data. Deep learning (DL) is gaining popularity due to its superior accuracy when learning large data volumes. Combining multiple ML or DL models is the primary component of hybrid approaches [53].

Table 1 provides a brief insight into various machine learning algorithms used in the analysis of environmental modelling. Here, it gives the various data taken from various stations installed all over the world and pollutants that are affecting human life.

Researchers have widely used CNN-LSTM architectures to forecast AQI, but several research gaps exist. Most studies concentrate more on accuracy as the primary evaluation criterion, with less emphasis on interpretability and feature importance. Only a few studies rigorously incorporate interpretability techniques such as SHAP, LIME, and PDP to investigate the contribution of individual features. Moreover, most studies base their work on extensive datasets and are not concerned with small-to-medium datasets or spatio-temporal heterogeneity among monitoring stations. Researchers seldom offer detailed analysis based on various measures of errors (MSE, RMSE, MAE) and statistical confirmation (one-way ANOVA), limiting the strength and faithfulness of the conclusions. This study fills the existing gaps by integrating CNN-LSTM prediction with interpretability analysis and detailed performance evaluation. Table 2 compares alternative methods (GRU, Transformers, GNNs, TabNet).

Table 2 emphasizes that although such techniques are effective, these approaches alternatively need a big dataset, insufficient in spatio-temporal incorporation, or are calculation-wise costly in comparison. CNN-LSTM offers a practical, precise, and reliable approach suited to the current data.

3 Methods

3.1 Study area

The Anugul region in Odisha was selected strategically as the study site for air quality forecasting because it has distinctive and urgent environmental issues that vary significantly from those of megacities. Unlike megacities, where on-road emissions are the primary source of pollution, Anugul suffers from severe deterioration in low air quality primarily due to large-scale industrial operations. The district is home to several big industries like coal-based power houses, aluminium smelters, and steel plants, and is among the most industrialized regions in eastern India [11]. The density of industries raises particulate matter and other contaminants, producing localized air quality issues that national-level studies cannot capture [40].

Anugul is not a focus of common air pollution research, mainly on large cities. Examining a non-megacity such as Anugul allows for investigating regional pollution dynamics and learning about industry-induced pollution trends. It is essential to understand

Table 1 Related work on machine learning tools application on air pollution modelling

Sl. No. and year	Author(s)	Dataset used	ML/DL algorithms applied	Pollutant(s) studied	Pre-processing	Performance metrics	Tool(s)/hardware employed	Key results
1 [63]	Beijing Multi-Site AQI Dataset	CNN-LSTM	PM _{2.5} , PM ₁₀ , NO ₂ , CO, SO ₂ , O ₃	Normalization, Sliding window, PCA	RMSE, MAE, R ²	TensorFlow, Python	Hybrid model outperformed standalone LSTM, RMSE reduced by 18%.	
2 [10]	Chinese AQI + Meteorological	ARIMA + CNN + LSTM (Optimized by DBO)	PM _{2.5} , PM ₁₀ , SO ₂ , NO ₂ , O ₃	Decomposition, Normalization	RMSE, MAE, R ²	Python, Scikit-learn	Hybrid model achieved RMSE = 4.13. DBO enhanced convergence rate.	
3 [13]	Turkey National AQI Data	CNN + LSTM	PM ₁₀ , SO ₂ , NO ₂	Feature scaling, Temporal encoding	RMSE, MAE	Keras, Python	CNN captured local variation; LSTM refined sequence learning.	
4 [31]	China National AQI + Climatic	CNN-LSTM-KAN Hybrid	PM _{2.5} , O ₃	Topography-aware input encoding	RMSE, MAE	TensorFlow, Python	PM _{2.5} affected most by topographic shifts; CNN localized spatial anomalies.	
5 [7]	OpenAQ + Meteorological	CNN-LSTM (Optimized by POA)	PM _{2.5} , NO ₂	Quantile encoding, Lag generation	RMSE, MAPE	PyTorch, POA-Optimizer	POA reduced RMSE by 17% compared to baseline.	
6 [47]	Hanoi AQI Data	CNN-LSTM + Attention + QPSO	PM ₁₀ , CO, NO ₂	Bayesian optimization, Time-shifted embedding	RMSE, MAE	Python, Keras	Quantum optimization improved interpretability and accuracy.	
7 [60]	Shijiazhuang AQI (2017–2021)	CNN + ILSTM (Improved LSTM variant)	AQI composite (PM _{2.5} -based)	CLM gating, deleted output gate, CNN feature extraction	MAE = 8.41, MSE = 202.19, R ² = 0.9601	Python, TensorFlow, Keras	Improved ILSTM accelerates training (~ 85 s) and boosts accuracy over CNN-LSTM baseline.	
8 [37]	Beijing PM _{2.5}	CNN + LSTM hybrid model	PM _{2.5}	Data windowing (7-day history), normalization	MAE, RMSE	Python, TensorFlow	The multivariate CNN-LSTM model outperformed univariate and multivariate LSTM baselines in MAE & RMSE.	
9 [28]	Dezhou City PM _{2.5} data set (2014–2024)	CNN, LSTM, CNN-LSTM, CNN-GRU-LSTM ANN	PM _{2.5}	Data normalization, feature scaling	MAE, RMSE, R ²	Python, Tensorflow	with faster training and superior prediction accuracy.	
10 [29]	Liaocheng city PM _{2.5} data (2014–2022)		PM _{2.5}	Data normalization, feature scaling	MAE, R ² , RMSE	MATLAB	CNN-GRU-LSTM outperformed other models in forecasting PM _{2.5} concentrations.	
								ANN model demonstrated high accuracy in forecasting monthly PM _{2.5} concentrations.

these trends to design practical, targeted mitigation efforts and inform policy-making for other similar semi-urban and industrial areas in India. Thus, this research fills a void in the current literature and adds to a more complete national understanding of air quality management. The Odisha Pollution Control Board (OPCB) official website (» Environmental Monitoring Data: State Pollution Control Board, Odisha (ospcboard.org)) offers the air quality data. This study used monthly average air quality data from five Anugul City monitoring stations.

This research utilizes the monthly average 2024 data from 44 monitoring locations in 17 Odisha cities and relates it to the Anugul data collected by the State Pollution Control Board (Odisha CPCB) website. The daily average data of 26 cities of India has also been used from 01-01-2015 to 01-07-2020, which is taken from the Central Pollution Control Board (CPCB) website, <https://cpcb.nic.in/National-Air-Quality-Index/>, in order to enhance the boundaries of the study area and to improve the data size and stations to get better results, as more data gives better results. The above Fig. 1 presents a map of Anugul, Odisha, and India.

3.2 Seasonal variations in AQI across data sets (city of Anugul, cities of Odisha, cities of India)

The bar graph shown in Fig. 2 compares the average Air Quality Index (AQI) over four seasons, Winter, Summer, Monsoon, and Post-Monsoon of five monitoring stations in the Anugul area.

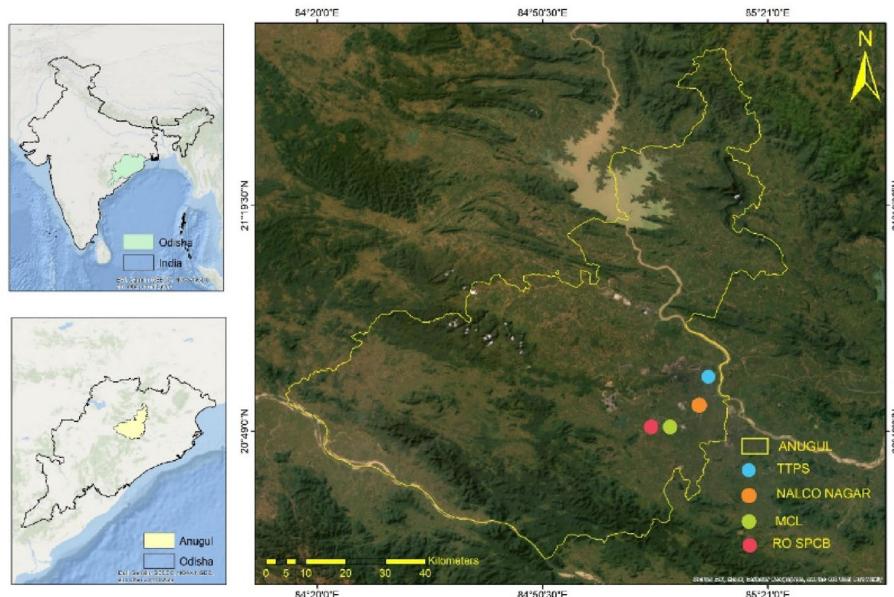
The graph shows that winter always has the highest AQI values at all stations, reflecting worse air quality during winter. Notably, the Mahanadi Coal Field Limited (MCL) and Talcher Thermal Power Station (TTPS) report relatively high AQI during winter, above 120, probably because of higher emissions and poor meteorological conditions. Conversely, the monsoon season has the minimum AQI values at all stations due to rain reducing suspended particulate matter in the air. Summer and post-monsoon seasons have moderate AQI values, an increase over monsoon but less than winter. MCL consistently reports the highest AQI readings of all stations in all seasons, whereas NALCO Nagar and RO SPCB have comparatively better air quality. These seasonal trends reflect the role of climate and industrial activity in air pollution, underpinning the necessity for season-specific mitigating measures.

A study in Jinan compared variations in the Air Quality Index (AQI) and the concentrations of the six primary air pollutants from 2014 to 2021. It emphasized that meteorologically unfavourable conditions were one of the significant factors behind serious winter pollution events [18].

The time series plot named “AQI Time Series per Station” in Fig. 3 shows the fluctuation in Air Quality Index (AQI) during January 2019 to December 2023 at five monitoring stations in the Anugul area. During this period, there has been a uniform increase in AQI values, which reflects progressive air quality degradation at all stations. The Mahanadi Coal Field Limited (MCL) station (green line) has persistently recorded higher AQI values, with some precipitous peaks, indicating recurrent spikes of likely mining and industrial origin. Similarly, RO SPCB and Talcher Thermal Power Station (TTPS) (both variants) also exhibit significant fluctuations and rising trends, indicating the role of industrial emissions. Conversely, NALCO Nagar (blue line) has comparatively stable and lower AQI values, meaning healthier air quality than other areas. The graph highlights

Table 2 Comparison table of alternative models

Model	Strengths	Limitations in present work context	Suitability for current study
CNN-LSTM [32]	Identifies spatial and temporal trends accurately, yields good results for limited to moderate datasets, Stable prediction performance and transparency, fast in terms of Computation	May not identify extremely prolonged correlations as accurately as transformers	Most suitable because of dataset size, efficiency, and explainability (SHAP, PDP)
GRU [14]	Less complex and quicker as compared to LSTM. Deals with time based relationships	A bit lower descriptive in relation to LSTM regarding extended series, comparable correctness as compared with LSTM but lacking positional CNN property retrieval	Can be employed, however gives lacking key benefit compared to CNN-LSTM
Transformers [2]	Outstanding in detecting extended connections, adaptable for big database	Needs significant learning inputs for consistency, algorithmically costly, explainability is very complicated	Non-appropriate in view of insufficient periodic AQI data
Graph Neural Networks (GNNs) [65]	Models structural site to site correlation is accurate, appropriate in case of location framework link is taken into account	Needs extensive locational graph data and massive samples, more difficult to explain finding, intensive calculation	Restricted application in this context as research concentrates upon sequential location-wise AQI instead of framework link
TabNet [41]	Deals with structured attributed efficiently, offers attribute significance clarity	Developed for grid-like categorization or trend analysis, instead of sequential temporal data unable to detect time-based connections immediately	Not perfect suited to AQI temporal data.

**Fig. 1** Map of Anugul, Odisha, and India

the necessity for improved air quality control, especially around high-emission industrial areas.

The graph in Fig. 4 presents seasonal trends in AQI per station, with the worst during winter, probably due to atmospheric inversion and concentration of pollutants. Monsoons exhibit the cleanest air, and summer and post-monsoons exhibit intermediate

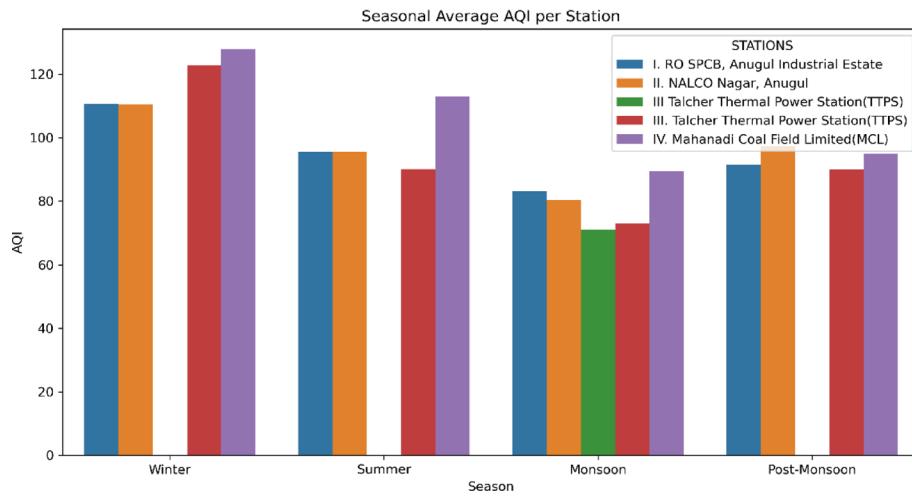


Fig. 2 Seasonal average AQI per station in Anugul Data

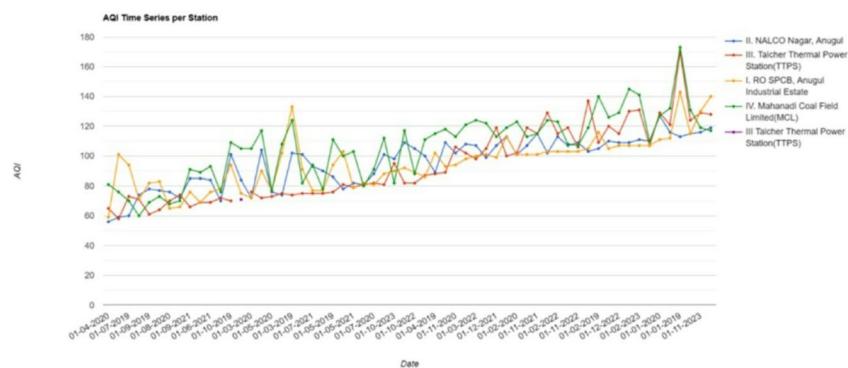


Fig. 3 AQI time series per station

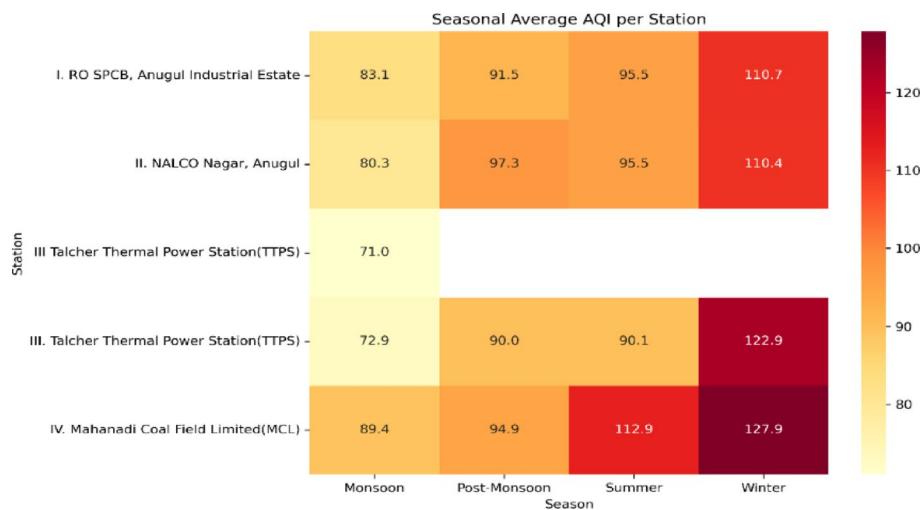


Fig. 4 Seasonal average AQI per station

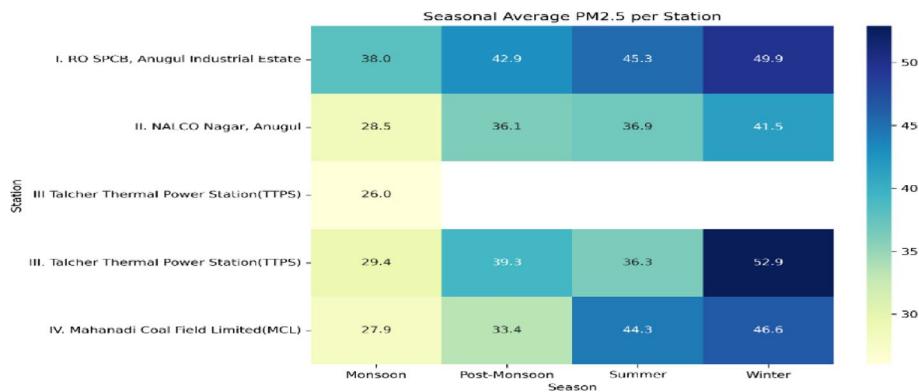


Fig. 5 Seasonal average PM_{2.5} per station

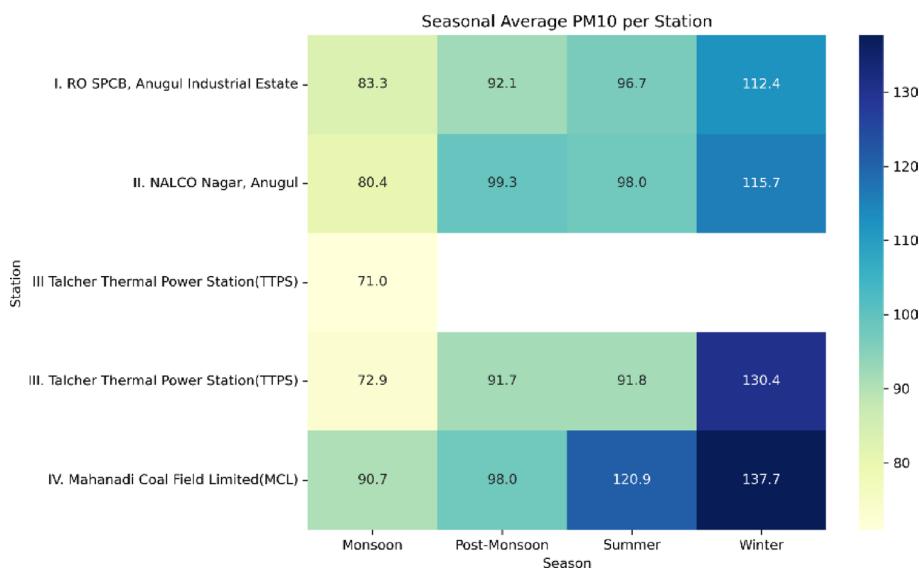


Fig. 6 Seasonal average PM₁₀ per station

pollution levels. These observations necessitate season-wise air quality management strategies, especially for winter periods.

Figure 5 above shows the seasonal variations in PM_{2.5} levels, with the winter period being the most polluted season and the monsoon the cleanest. The data point to the necessity of season-specific pollution control measures, especially for particulate matter during the winter season.

The seasonal average PM₁₀ concentrations at five monitoring stations in Anugul, Odisha, are shown in the above heatmap (see Fig. 6). The Mahanadi Coal Field Limited (MCL) station recorded the peak value of 137.7 µg/m³, indicating that PM₁₀ levels are typically highest during the winter. On the other hand, PM₁₀ concentrations are lowest during monsoon seasons, especially at the Talcher Thermal Power Station (TTPS), where they are 71.0 µg/m³. The trend points to a seasonal change in air pollution, with the air quality being noticeably worse in the winter. Figures 4 and 5, and 6 present the AQI, PM₁₀, and PM_{2.5} seasonal trends. The results suggest that the behaviour of most pollutants has a specific seasonal representation. This indicates the importance of the environmental policy in the different seasons and continuous air quality monitoring.

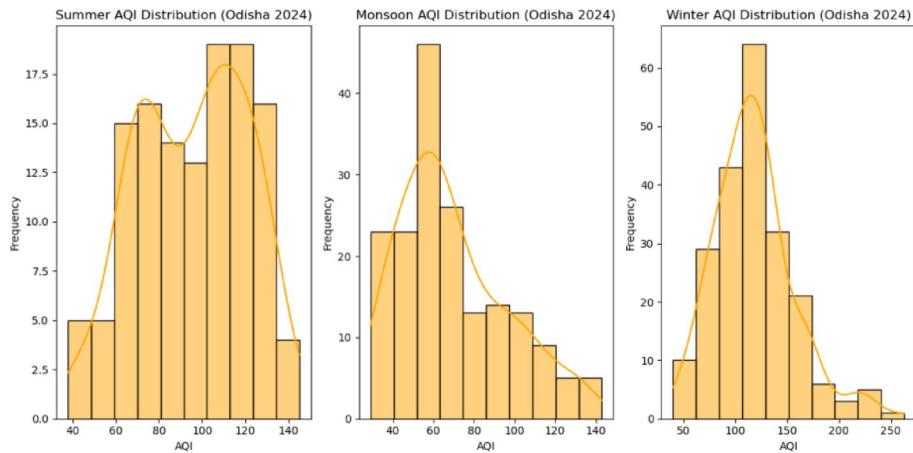


Fig. 7 Seasonal average AQI per station in Odisha data

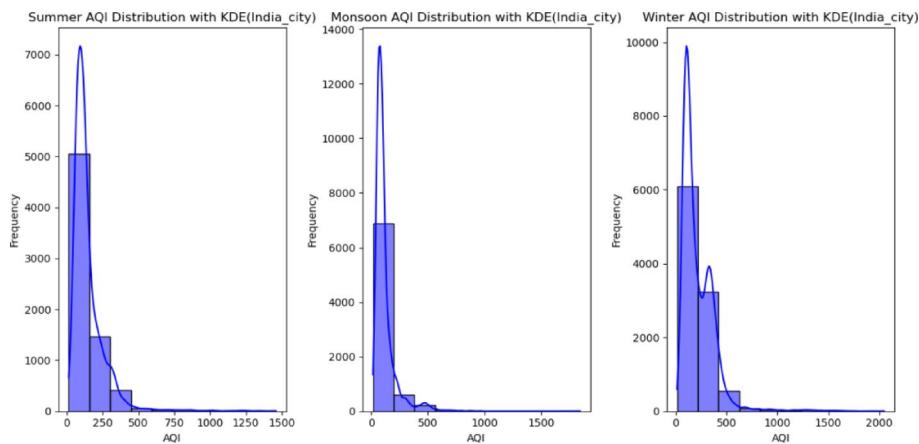


Fig. 8 Seasonal average AQI per station in cities in India data

Figure 7 represents the Seasonal distributions of AQI over monitoring stations in Odisha during 2024, which exhibit clear temporal trends. During the summer season, AQI values are moderately spread over the range of 60–120, indicating sources from dust resuspension and traffic activities. During the monsoon season, the distribution shows a definite leftward bias with most stations reporting AQI in the range of 50–80, implying efficient removal of atmospheric pollutants through wet deposition. Conversely, winter shows a marked rightward shift with AQI values ranging from 100 to 150 and extending well beyond 200, reflecting worsened air quality. Stable atmospheric conditions, low boundary layer height, and seasonal biomass burning are the causes of this deterioration. These station-wise distributions highlight the dominant role of seasonal meteorological conditions in regional air quality dynamics.

The Seasonal distributions of AQI over Indian cities depict apparent variability during summer, monsoon, and winter (See Fig. 8). The distribution for the summer season depicts values of AQI mostly concentrated below 200 with sparse occurrences of high pollution levels, which suggests moderate air quality conditions. Under the monsoon season, the distribution shifts further to the left with most AQI values below 150, reflecting the beneficial impact of rain in diluting ambient pollutant concentrations. Conversely, the winter season reveals a rightward shift in the distribution, with increased

frequencies of AQI values in the range 200–400, indicating significant deterioration in air quality induced by stable atmospheric conditions and seasonal emission sources. These distributions reflect the role of meteorological factors and seasonal activities in influencing air quality over Indian cities.

3.3 SHAP visualization on pollutants

SHAP (Shapley additive explanations) values are a powerful transparency tool for explaining the machine learning model training. SHAP values find the influence of each variable by calculating all possible combinations of variables [59]. Users assimilate the overall importance of the factors, as well as the impact of each factor on a particular prediction. It benefits complex models like ensembles and deep learning [16]. The SHAP value is shown below in Fig. 9.

The summary plot shows the SHAP value for how each feature affects the model's decisions. For a feature and a data point, each dot represents a single SHAP value. The x-axis is the SHAP value, which represents the contribution of a feature to increasing or reducing the model output [1]. The SHAP value dispersion highlights each feature's significance, such as NO_2 , O_3 , SO_2 , NH_3 , $\text{PM}_{2.5}$, and PM_{10} , displayed on the y-axis. Each dot's colour represents the feature's real value; red denotes a high feature and blue a low one [35]. With high values (red) pushing the forecast in a favourable direction, PM_{10} , for instance, has the broadest range of SHAP values, suggesting it has the most significant influence on the model's predictions. Low PM_{10} concentrations (blue) reduce the model output. Both the significance of the features and the direction of their impact on the model prediction are revealed by the above graph, shown in Fig. 9.

The LIME explanation for Anugul (see Fig. 9) reveals that $\text{PM}_{2.5}$ and NH_3 have positively contributed to raising the calculated AQI, whereas PM_{10} , NO_2 , SO_2 , and O_3 have an adverse effect. The overall effect of these pollutants provides a locally predicted AQI value of approximately 101 (Fig. 10).

Figure 11 displays the SHAP summary plot for the Odisha data, showing that PM_{10} and $\text{PM}_{2.5}$ predominantly influence AQI prediction, and high concentrations of these pollutants (in red) strongly increase the AQI. In contrast, low concentrations (in blue) decrease the AQI. In the gaseous pollutants, NO_2 , SO_2 , O_3 , and NH_3 have relatively small influences, suggesting that they have a limited contribution to AQI dynamics in the area.

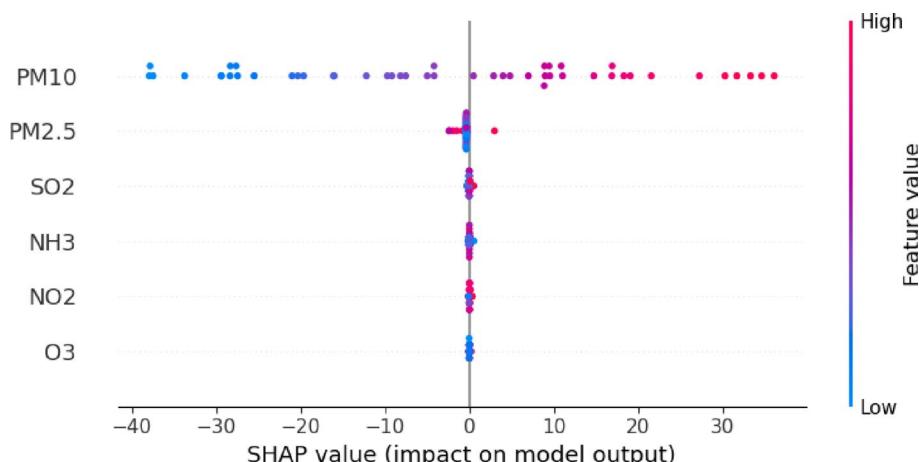


Fig. 9 SHAP visualization of pollutants of Anugul

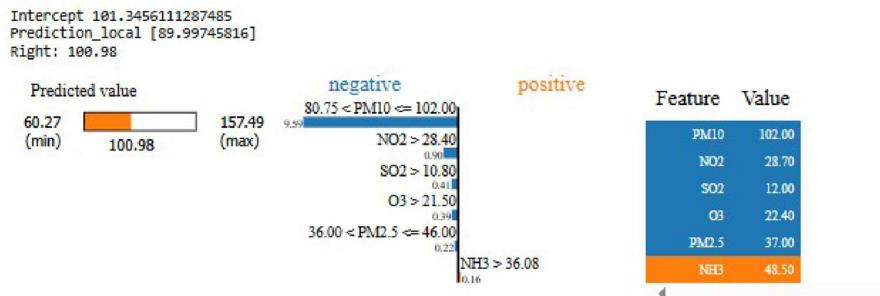


Fig. 10 LIME visualization of pollutants of Anugul

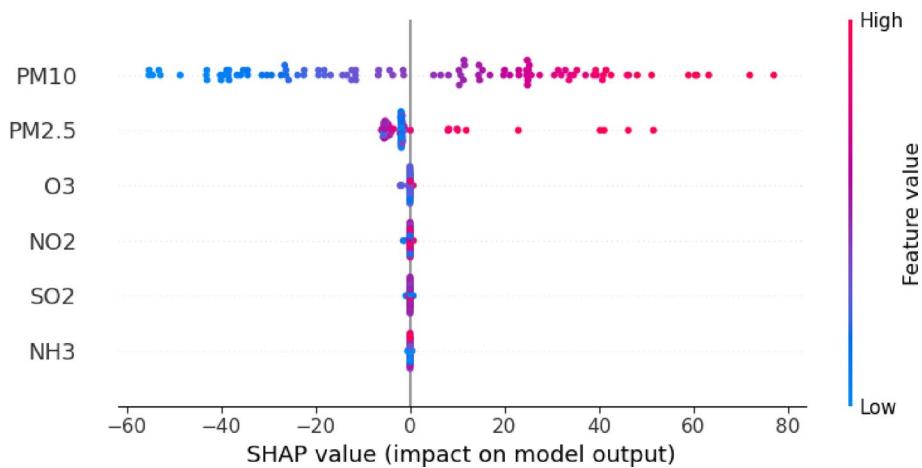


Fig. 11 SHAP visualization of pollutants of Odisha

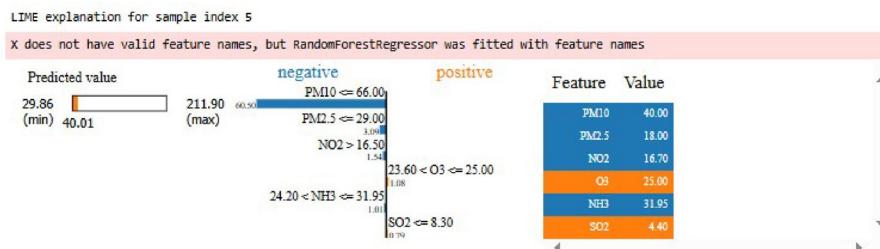


Fig. 12 LIME visualization of pollutants of Odisha

The extensive range of SHAP values for PM_{10} , varying from strongly negative to highly positive, indicates that changes in its concentration significantly impact the prediction results. $\text{PM}_{2.5}$ also shows a well-defined positive correlation with AQI, supporting its role in air quality episode determination. The findings verify that particulate matter, particularly PM_{10} and $\text{PM}_{2.5}$, is the decisive parameter influencing AQI levels in Odisha, with less prominent contributions from gaseous pollutants.

The LIME explanation of Odisha data (see Fig. 12) reveals that PM_{10} , $\text{PM}_{2.5}$, NO_2 , and NH_3 reduce the prediction of AQI negatively, whereas O_3 and SO_2 enhance it. The sum effect of these air pollutants leads to a locally predicted AQI of around 40.

In the India dataset, the SHAP values (see Fig. 13) show that $\text{PM}_{2.5}$ and PM_{10} are the most prominent factors influencing AQI, with increased concentrations (red)

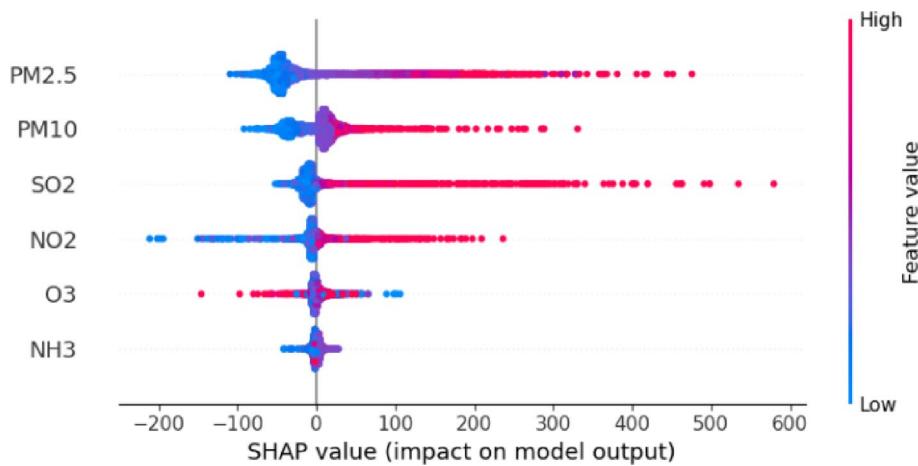


Fig. 13 SHAP visualization of pollutants of India

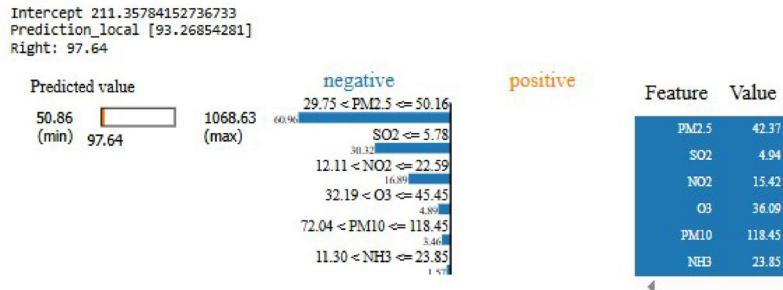


Fig. 14 LIME visualization of pollutants of India

contributing substantially to the model prediction and, in turn, deteriorating the air quality. The size of the SHAP values for $\text{PM}_{2.5}$ exceeds 500, indicating its leading contribution to influencing AQI predictions nationwide. SO_2 and NO_2 also indicate a significant positive effect, especially at elevated levels, identifying their secondary contribution to pollution events. O_3 and NH_3 have comparatively minor effects, with their SHAP values closer to zero, indicating they contribute less towards determining AQI. The findings highlight that particulate matter is the key factor determining air quality in India, and that gaseous pollutants such as SO_2 and NO_2 have additional but secondary effects.

The LIME explanation of India data (See Fig. 14) shows that $\text{PM}_{2.5}$, SO_2 , NO_2 , O_3 , PM_{10} , and NH_3 all have a negative contribution, thus reducing the predicted AQI. Their net impact leads to a locally predicted AQI of approximately 98.

Although SHAP is a powerful interpretability tool, it has its own set of limitations. Firstly, it relies on feature independence and additivity assumptions in splitting contributions, which might not always be the case for highly correlated environmental variables such as air pollutants. Secondly, the computational expense can be extremely high for large datasets or complex models, rendering it challenging to use in real-time or large-scale applications. Thirdly, SHAP values are sometimes unstable when the data is imbalanced and can result in incorrect interpretations if not thoroughly cross-checked [58].

Despite these drawbacks, SHAP is still one of the most used interpretability techniques in air pollution prediction research. It helps determine the primary contaminants affecting AQI since it may offer reliable, model-agnostic answers. Consequently, academics

widely use SHAP to improve trust and transparency in deep learning and hybrid models used with environmental data.

4 Key architecture of predictive models

4.1 Traditional model for time series forecasting

ARIMA and Seasonal ARIMA (SARIMA) models [57] are examples of time series forecasting. The ARIMA model addresses the effects of random shocks and the autocorrelation in stationary time series, a basic technique for modelling this data type [42]. An extension of ARIMA is called SARIMA, which was created primarily to handle time series that show seasonal trends [34].

When building ARIMA and SARIMA models, the correct parameters are needed. The ideal parameters can be obtained within a specified search range using metrics like the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC) [6]. With little external control and big, regular datasets, these conventional time series models typically perform well for short-term forecasting. However, their predictive ability significantly declines when the external variables are highly changeable [38].

P. Hasan et al. [27] validated the performance of the ARIMA and SARIMA models with traditional measures of accuracy. For the ARIMA (1,1,1) model, the given results were MAE = 1266.08, MSE = 28,550,102.047, and RMSE = 5343.230, whereas for the SARIMA (1,1,2) (1,1,2)12model, the given results were MAE = 1246.442, MSE = 26,789,531.837, and RMSE = 5175.860, showing good predictive ability of both models.

4.2 RNN

Dai et al. [8] characterize a Recurrent Neural Network (RNN) as having full connectivity between consecutive layers and between neurons in the same layer. With this structure, a neuron's output can be used as input for another, thus allowing information to pass between neurons. Figure 15(a) shows the structure of the RNN. RNNs do not capture long-range dependencies, although they are effective at learning short-range connections in time series data.

4.3 LSTM

Gong et al. [15] demonstrate that the Long Short-Term Memory (LSTM) neural network solves the problem of long-distance dependence very well. According to Salim and Djunaidy [52], it comprises an input layer, an LSTM layer, a fully connected layer, and

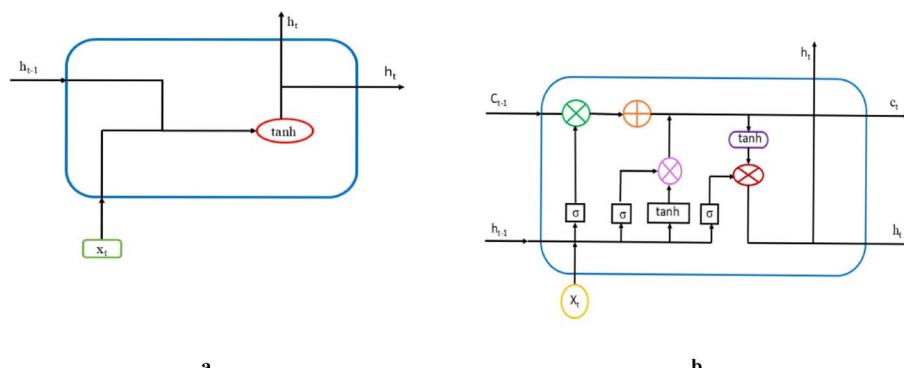


Fig. 15 **a** RNN neuron memory cell. **b** Memory cell of LSTM neuron

an output layer. LSTM comprises input, forget, and output gates to control data flow, as shown in Fig. 15(b). These gates regulate how data is kept, handled, or shared. It overcomes the vanishing and expanding gradients and provides important information over lengthy sequences.

4.4 Convolutional neural network (CNN)

CNN is a collection of layers taking some weights that enable extracting some features in a specific area or location [55]. The number of training parameters is adjusted and enhances the computational complexity in the model to remove the possibility of overfitting [46]. Figure 16 shows the CNN architecture.

4.5 The proposed CNN-LSTM model

In this research, a CNN-LSTM model integrates CNN and LSTM to enhance the efficacy of AQI prediction. The input layer takes historical information regarding air quality and related environmental factors. These typically consist of pollutant levels ($\text{PM}_{2.5}$, PM_{10} , NO_2 , SO_2 , CO , O_3), meteorological conditions (temperature, humidity, wind speed), and timestamps. Following feature extraction, a one-dimensional vector is created by flattening the output. The data is ready for the LSTM layers to process sequentially after this phase. LSTM units are appropriate for time-series forecasting applications like AQI prediction because they efficiently learn from historical observations using memory cells and gating mechanisms (input, forget, and output gates). The final output is usually produced by adding dense layers and performing additional non-linear operations. Figure 17 illustrates the model.

Figure 17 is a hybrid deep learning model to predict the Air Quality Index (AQI) from the inputs of various pollutants. Certain air pollutants, including SO_2 , NO_2 , and PM_{10} , are used as input. A convolutional neural network (CNN) learns spatial information about the inputs taken. Each CNN output is shifted to a corresponding LSTM network to produce the long-term trends and temporal relationships. Finally, the output layer generates AQI predictions from all three LSTMs, such as AQI_1 , AQI_2 , and AQI_3 , which are the AQI values for each pollutant in a given station.

5 Simulation

5.1 Experiment set up

The experiment employed Python 3.9 and open-source modules such as Keras, NumPy, and Pandas. TensorFlow is installed in the Anaconda Navigator environment to carry out the training process. The computer runs the experiment on a 64-bit Windows 11,

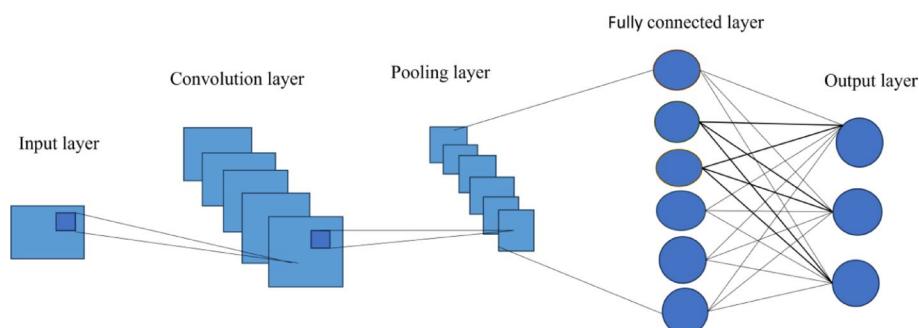


Fig. 16 Architecture of CNN

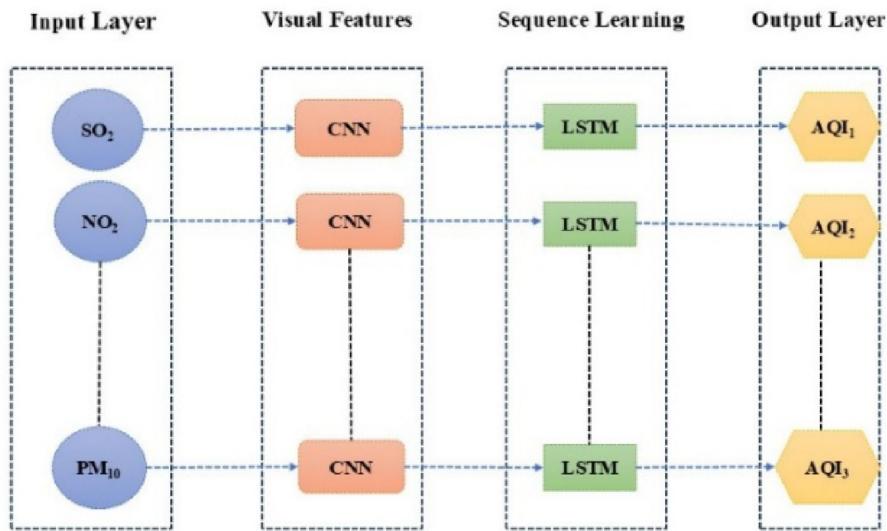


Fig. 17 CNN-LSTM architecture

2.10 GHz processor, and 8 GB of RAM. The performance of models like CNN, LSTM, and CNN-LSTM was evaluated using statistical criteria, which were Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). In every model, the Air Quality Index (AQI) is predicted based on the air pollutants such as SO₂, NO₂, NH₃, O₃, PM₁₀, and PM_{2.5} taken as input features. Figure 18 is a flowchart of the training process in the CNN-LSTM network to assess the AQI index.

The model is pre-processed with the input features and normalized before the training distribution. The process divides the dataset into test and training sets. The trained model predicts the AQI values using test inputs and compares them with actual values to calculate the parameters that determine the model's accuracy and performance. This process provides temporal dependencies and informative sequences within the dataset. The various layers within the CNN-LSTM network for predicting the AQI index are provided below in Fig. 19.

For CNN-LSTM-based prediction of AQI in Anugul, Odisha, and Indian city datasets, preprocessing begins with the selection of appropriate features (emissions like SO₂, NO₂, NH₃, O₃, PM₁₀, PM_{2.5}) and target AQI, and omitting missing or non-numeric data. Replaces any value with "BDL" (Below Detection Limit) by 0, transforming all data columns into numeric types. It then creates temporal sequences from normalized values, making predictions using a fixed-size window of previous observations (e.g., 5–30 days) for the next time step's AQI while retaining sequential information. The sequences are divided into training and testing (or validation) sets, maintaining the chronological sequence to avoid data leakage. The CNN-LSTM learns patterns from the data at model training time using backpropagation. The prediction error is propagated backwards through the convolutional, pooling, and LSTM layers, and the optimizer adjusts the network weights to minimize loss. Following training, predictions are reversed back to the original AQI scale, and model performance is tested using metrics including MSE, RMSE, and MAE, with plots of predicted vs. actual AQI serving as a visual check of predictive accuracy.

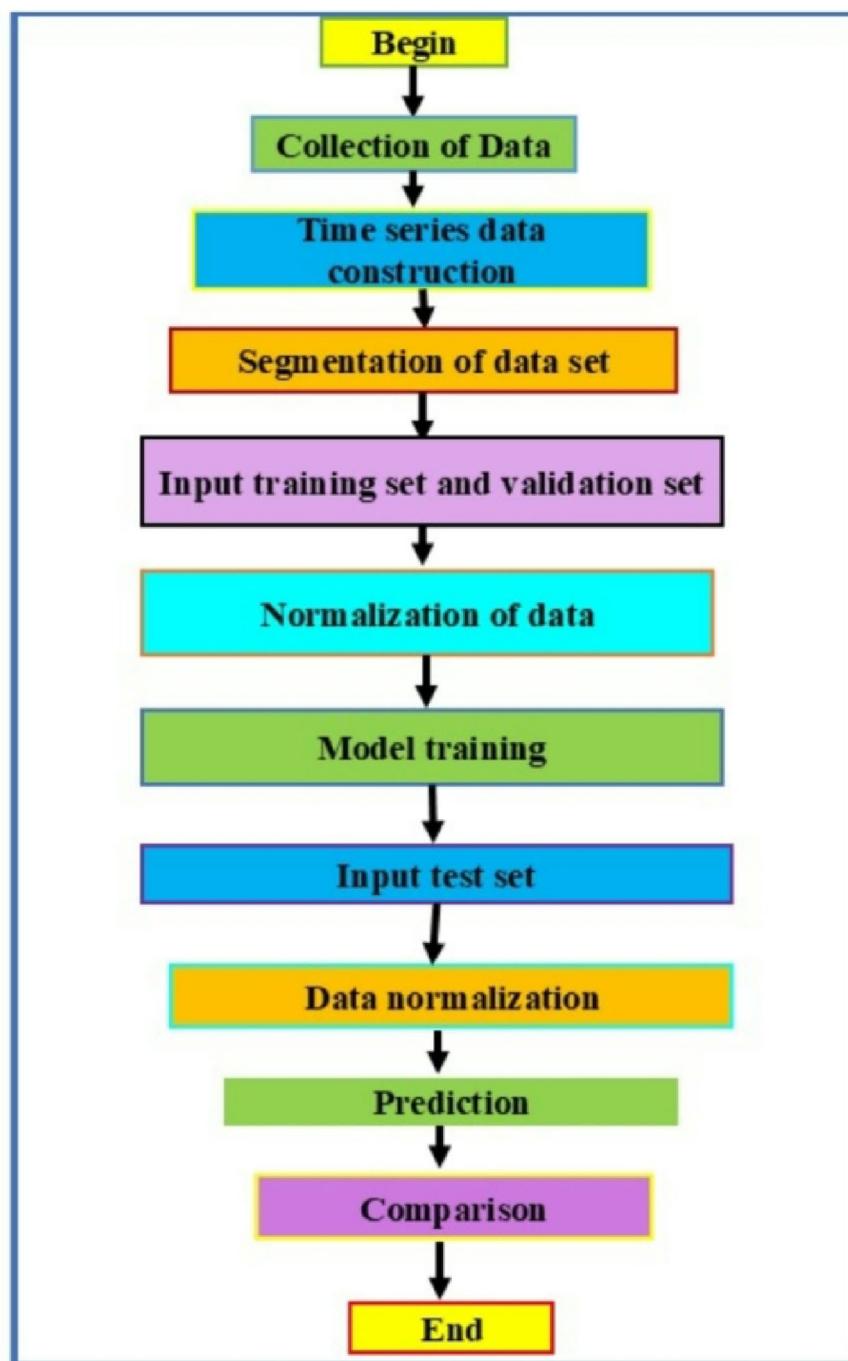


Fig. 18 Flowchart of experimental process

5.2 Impact of air quality parameters on AQI prediction using partial dependence plots

Partial Dependence Plots (PDPs) are tools to interpret the individual feature effects in the training process that evaluates the model output [51]. In this section, PDP analysis provides the most important air quality pollutants like SO_2 , NO_2 , NH_3 , O_3 , PM_{10} , and $\text{PM}_{2.5}$, and their related contributions to the Air Quality Index prediction. Figures 13, 14 shows the following:

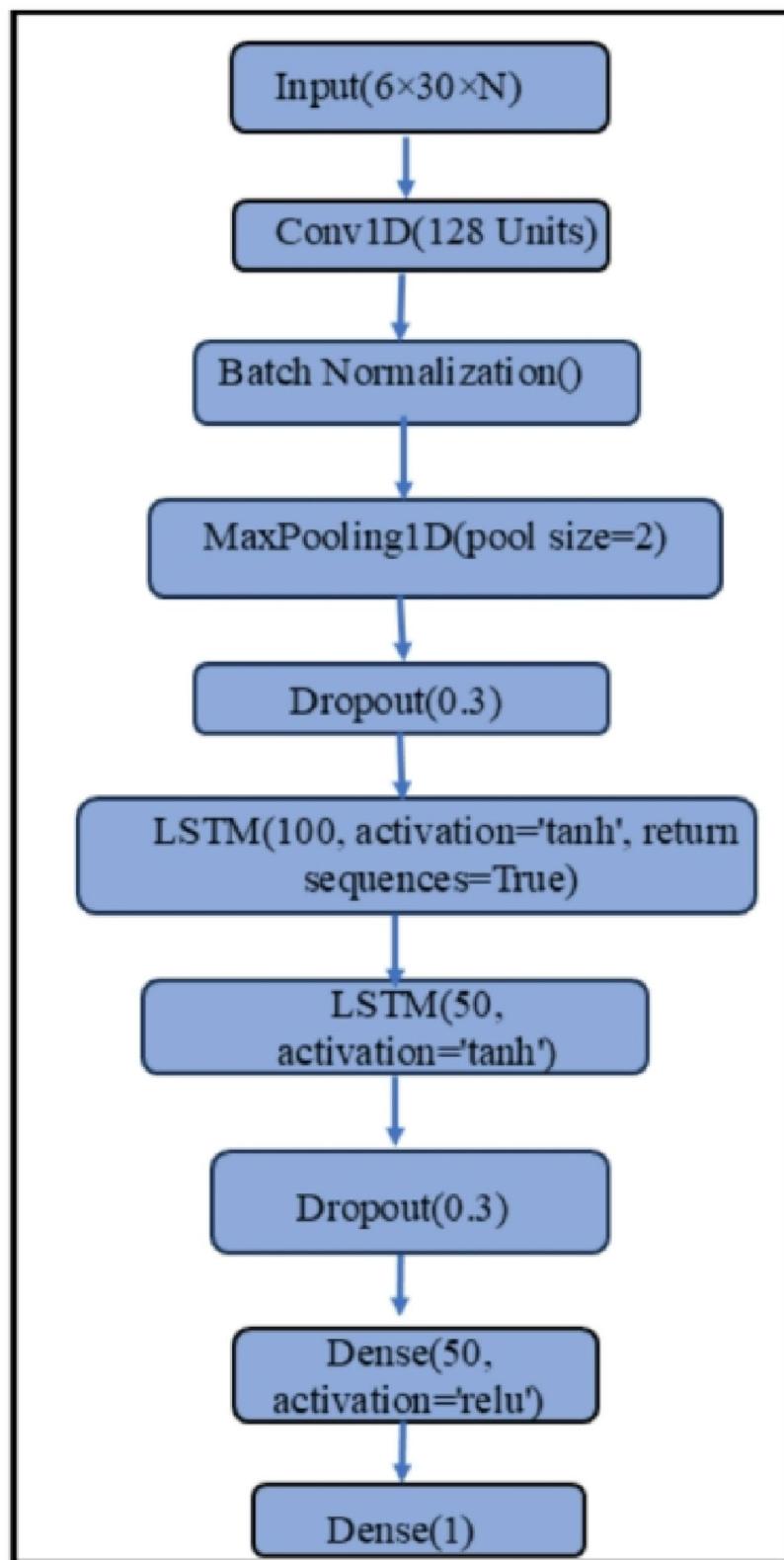


Fig. 19 Flowchart of simulation process in CNN-LSTM network for prediction of AQI index

- This trend is positively related to concentrations of SO_2 and AQI, showing that increased concentrations of SO_2 cause higher air pollution levels.
- Likewise, slight increments of NH_3 can enhance air quality; further increments have negligible effects on AQI.
- The PDP for O_3 shows oscillatory behaviour, which mirrors the intricate nature of ozone pollution.
- A distinct linear relationship exists between PM_{10} and AQI, where higher PM_{10} concentrations lead to higher AQI values. This result makes PM_{10} a significant determinant of air quality decline. The PDP for $\text{PM}_{2.5}$ indicates a rapid increase in AQI at higher concentrations. $\text{PM}_{2.5}$ is a dominant predictor of air quality.
- The PDP analysis reveals that $\text{PM}_{2.5}$ and PM_{10} are the dominant pollutants in the prediction of AQI, followed by NO_2 and SO_2 .

With the same preprocessing techniques described above for Odisha data, we developed Partial Dependence Plots (PDPs) for AQI. Following preprocessing, we trained a Random Forest Regressor model over the pollutant features to predict AQI and created PDPs using 'Partial Dependence Display'. The estimator calculates the average predicted AQI by changing each of the chosen features while setting the remaining ones to their observed values. This method separates the marginal effect of each pollutant on AQI and gives a neat visualization of how alterations in individual pollutant levels affect the prediction from the model.

The Partial Dependence Plots (PDPs) of Anugul (see Fig. 20) demonstrate the impact of varied pollutants on AQI predictions. From the plots, it can be observed that SO_2 , NO_2 , NH_3 , and O_3 have almost flat lines, reflecting negligible contributions to AQI predictions. On the other hand, PM_{10} shows a significant positive correlation with AQI, reaffirming it as the major driver of air pollution in this area. $\text{PM}_{2.5}$ is also increasing slightly, indicating a moderate but weaker contribution than PM_{10} . Generally, the analysis points out that particulate matter, particularly PM_{10} , has most significantly influence determining AQI levels, with gaseous pollutants having minimal contributions.

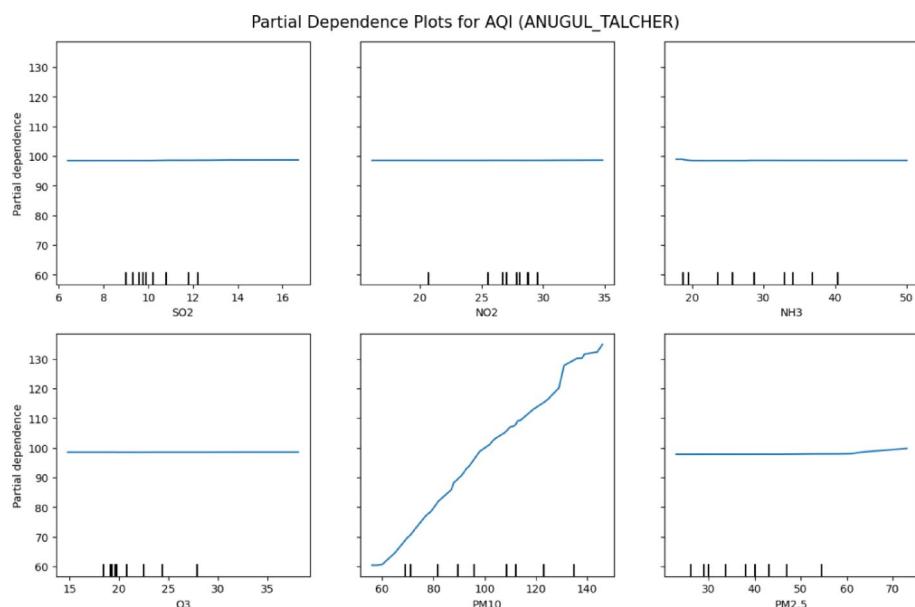


Fig. 20 PDP analysis of input parameters of Anugul

Partial Dependence Plots (PDPs) of Odisha (see Fig. 21) show that particulate matter, and especially PM₁₀, is the overarching parameter determining AQI, with a robust linear increase as concentrations rise, with PM_{2.5} also playing an important role, especially above higher cut-offs. NO₂ and SO₂ have almost flat trends, meaning minimal contribution to air quality in this area. These results emphasize that Odisha's degraded air quality is primarily caused by particulate pollution, whereas gaseous pollutants are unimportant. Hence, management policies for air quality should focus on controlling PM₁₀ and PM_{2.5} emissions from sources like industrial activity, road dust, construction, and coal burning.

Indian city PDP plots (see Fig. 22) show that PM_{2.5} is the controlling variable in influencing AQI, with concentrations increasing steeply and exceeding 300 at very high levels, whereas PM₁₀ increases consistently but with less intensity. NO₂ follows a weak rising trend, indicating mild influence, and SO₂ follows moderate impacts with episodic spikes, indicating localized effects. PM_{2.5} and PM₁₀ are the most dominant contributors to poor air quality, and gaseous pollutants are responsible to a lesser extent. These results suggest that strong control of fine and coarse particulate emissions is essential for better air quality in Indian cities.

Although partial dependence plots (PDPs) help understand the marginal contribution of single features towards model predictions, they are not without limitations. PDPs rely on independent features, which may not always be the case in real-world data where pollutants tend to be correlated with each other, and could thus lead to incorrect interpretations. Further, calculating PDPs can be computationally expensive in the case of larger datasets or more complex models such as Random Forests. Nevertheless, PDPs are still popular for model interpretability in environmental modeling [44].

In this experiment, three optimizers, Adam, RMSprop, and SGD, were used to collect data for Anugul city. The Adam optimizer serves as a type of first-order optimization

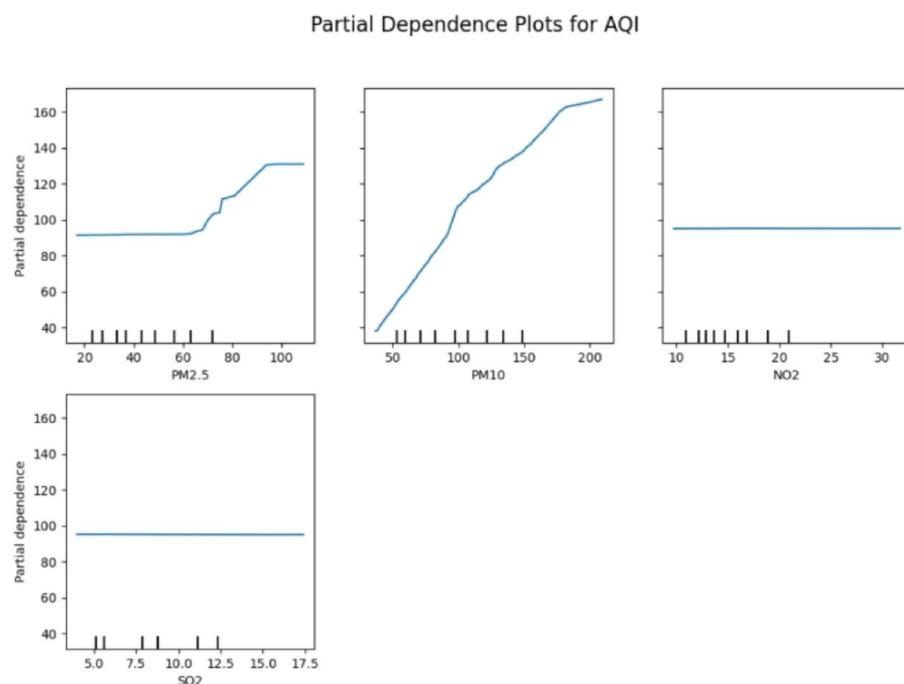


Fig. 21 Analysis of input parameters of Odisha

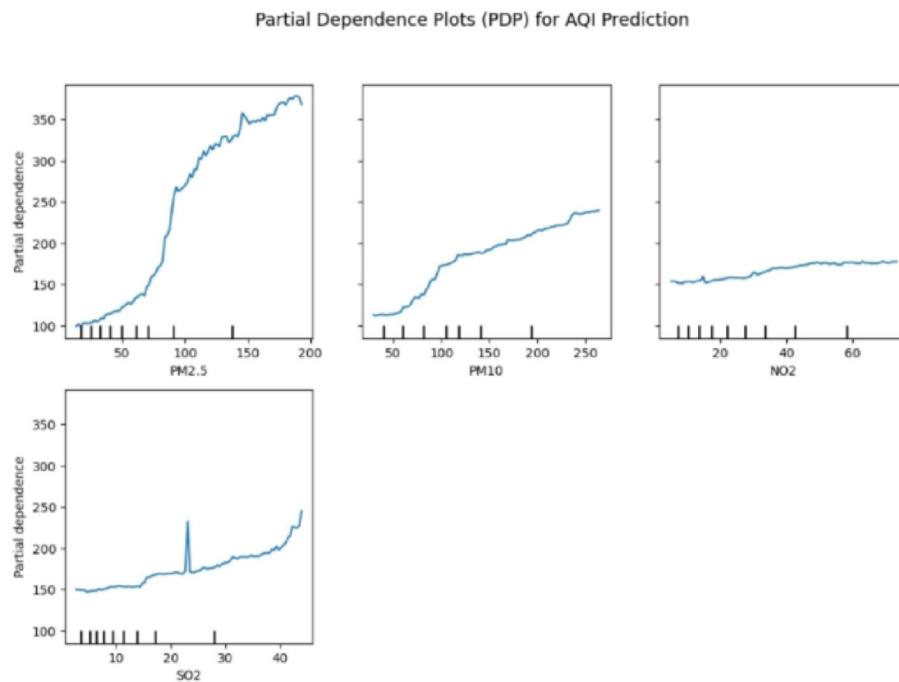


Fig. 22 Analysis of input parameters of India

Table 3 Hyperparameter tuning for Anugul data set

Hyperparameters	Values for Anugul data
Window Size	30
Batch Size	16
Learning Rate	0.001
Optimizer	Adam, RMSprop, SGD
CNN Filters	128
Kernel Size	5
LSTM Units	128
Dropout Rate	0.3
Dense Units	128
Activation Function	tanh
Epochs	100, 200, 300
Loss Function	MSE
Validation Split	0.2

algorithm. This optimizer functions to adjust the parameters in deep learning models [23]. Table 2 below provides the parameters and their values for these three optimizers.

Tables 3 and 4 shows the values for hyperparameters for Anugul, Odisha and India respectively.

The training and validation loss graphs of Adam, RMSprop, and SGD optimizers at 100, 200, and 300 Epochs are shown in Fig. 23(a)-(i) for Anugul data.

The plots contrast training and validation loss patterns for various optimizers (Adam, RMSprop, and SGD) over 100, 200, and 300 epochs. In all instances, the training loss drops steadily as epochs rise, which indicates that the models are learning well. The validation loss also follows the same trend but with occasional fluctuations, demonstrating the model's generalization ability on unseen data. Adam converges faster with less loss, whereas RMSprop stabilizes slowly and SGD minimizes more slowly. Overall, the

Table 4 Hyperparameter tuning for Odisha and Indian cities data set

Hyperparameters	Values for Odisha data	Values for Indian city Data
Sequence Length	5	30
Conv1D Filters	32, 64, 128	64
Conv1D Kernel Size	2, 3, 5	3
MaxPooling1DPool Size	2, 3	2
LSTM Units	64	64
Dropout Rate	0.3	0.2
Dense Layer Units	32	32
Batch Size	32	32
Learning Rate	0.001	0.001
Epochs	200, 300, 400, 500	20, 40, 60, 80

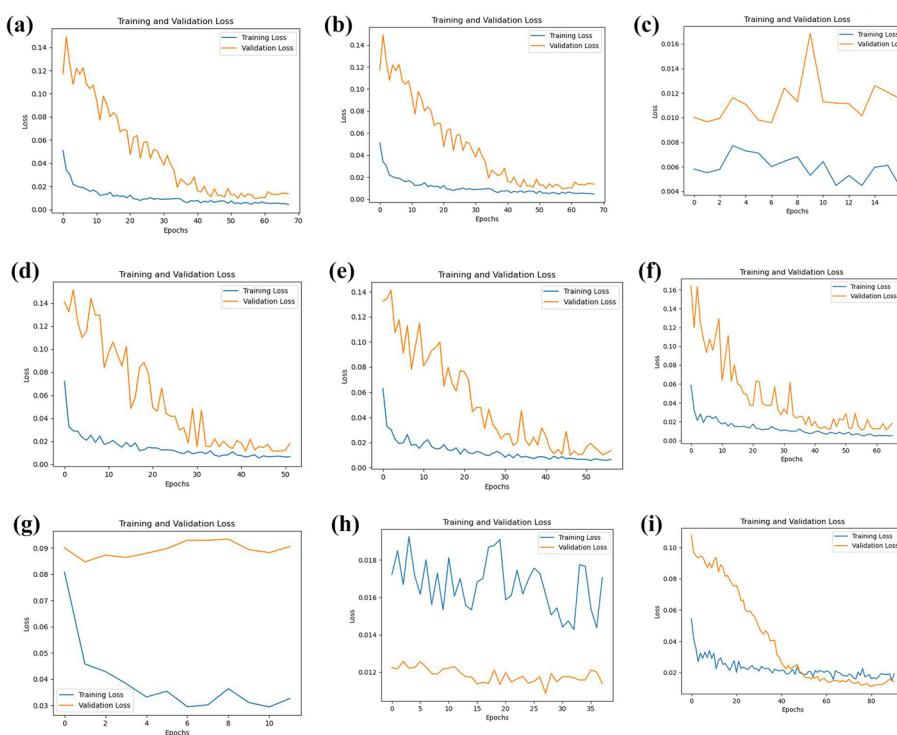


Fig. 23 **a** Training and validation loss graph of Adam optimizer at 100 epochs. **b** Training and validation loss graph of Adam optimizer at 200 epochs. **c** Training and validation loss graph of Adam optimizer at 300 epochs. **d** Training and validation loss graph of RMSprop optimizer at 100 epochs. **e** Training and validation loss graph of RMSprop optimizer at 200 epochs. **f** Training and validation loss graph of RMSprop optimizer at 300 epochs. **g** Training and validation loss graph of SGD optimizer at 100 epochs. **h** Training and validation loss graph of SGD optimizer at 200 epochs. **i** Training and validation loss graph of SGD optimizer at 300 epochs

plots emphasize how the choice of optimizer and training time heavily influence model performance.

The actual versus predicted graph of Adam, RMSprop and SGD optimizer at 100, 200 and 300 epochs are shown in Fig. 24(a)-(i) for Anugul data.

The plots of Anugul indicate the comparison of actual and predicted values via various optimizers (Adam, RMSprop, and SGD) with different epochs (100, 200, and 300). The prediction accuracy of the Adam optimizer enhances as epochs rise, with better matching of actual and predicted values at 300 epochs. Likewise, RMSprop indicates incremental improvement in prediction accuracy with greater epochs, though its performance treads more slowly than Adam's. The SGD optimizer begins with bigger

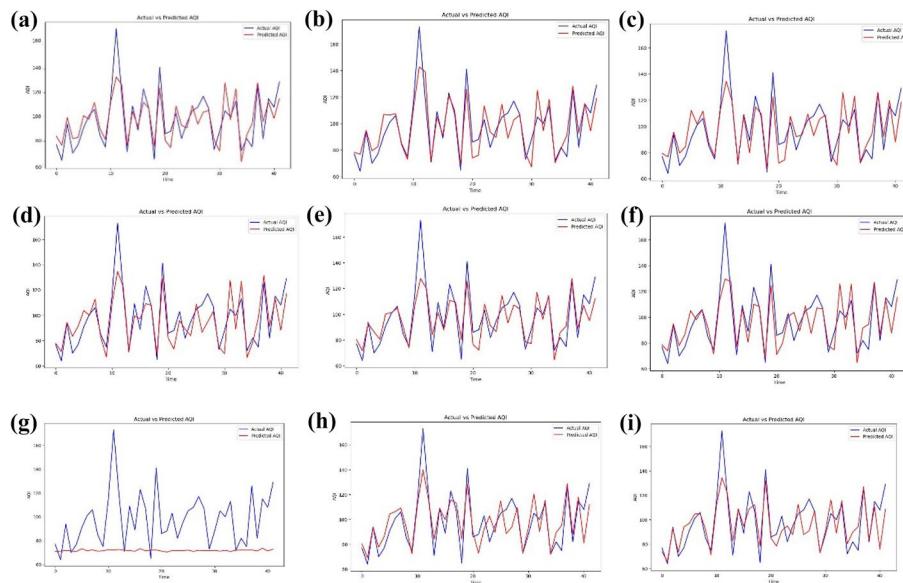


Fig. 24 **a** Actual and predicted graph of Adam optimizer at 100 epochs. **b** Actual and predicted graph of Adam optimizer at 200 epochs. **c** Actual and predicted graph of Adam optimizer at 300 epochs. **d** Actual and predicted graph of RMSprop optimizer at 100 epochs. **e** Actual and predicted graph of RMSprop optimizer at 200 epochs. **f** Actual and predicted graph of RMSprop optimizer at 300 epochs. **g** Actual and predicted graph of SGD optimizer at 100 epochs. **h** Actual and predicted graph of SGD optimizer at 200 epochs. **i** Actual and predicted graph of SGD optimizer at 300 epochs

deviations at 100 epochs but improves dramatically by 300, reflecting improved convergence with increased training. In summary, the graphs confirm that epoch increase improves prediction accuracy, with Adam doing the best to converge fast and stably compared to RMSprop and SGD.

The above graphs from Fig. 25a-g show the performance of the Adam optimizer at various training epochs (200, 300, 400, and 500) for Odisha's dataset. For every case, the graphs depict two: one for comparing the actual and predicted values, and the other for training and validation loss curves. As the number of epochs increases, the predicted values are closer to the actual values, enhancing the model's accuracy.

In the same way, the loss plots also indicate a slow reduction in training and validation loss with increased epochs, indicating improved convergence of the model. Generally, the plots illustrate how longer training improves the optimizer to reduce error and make predictions that more closely resemble the actual outcomes.

Figure 26a-d illustrates the predicted versus actual values for Indian cities using the Adam optimizer across multiple epochs (20, 40, 60, and 80). The predictions are noticeably away from the actual values at 20 epochs, showing underfitting. When the epochs rise to 40 and 60, the predicted values closely represent the actual data, demonstrating stronger learning and higher accuracy. By 80 epochs, the model shows good convergence, with predictions closely tracking the actual trend, indicating that the optimizer has successfully reduced errors. Generally, the graphs indicate improving prediction accuracy with additional training epochs.

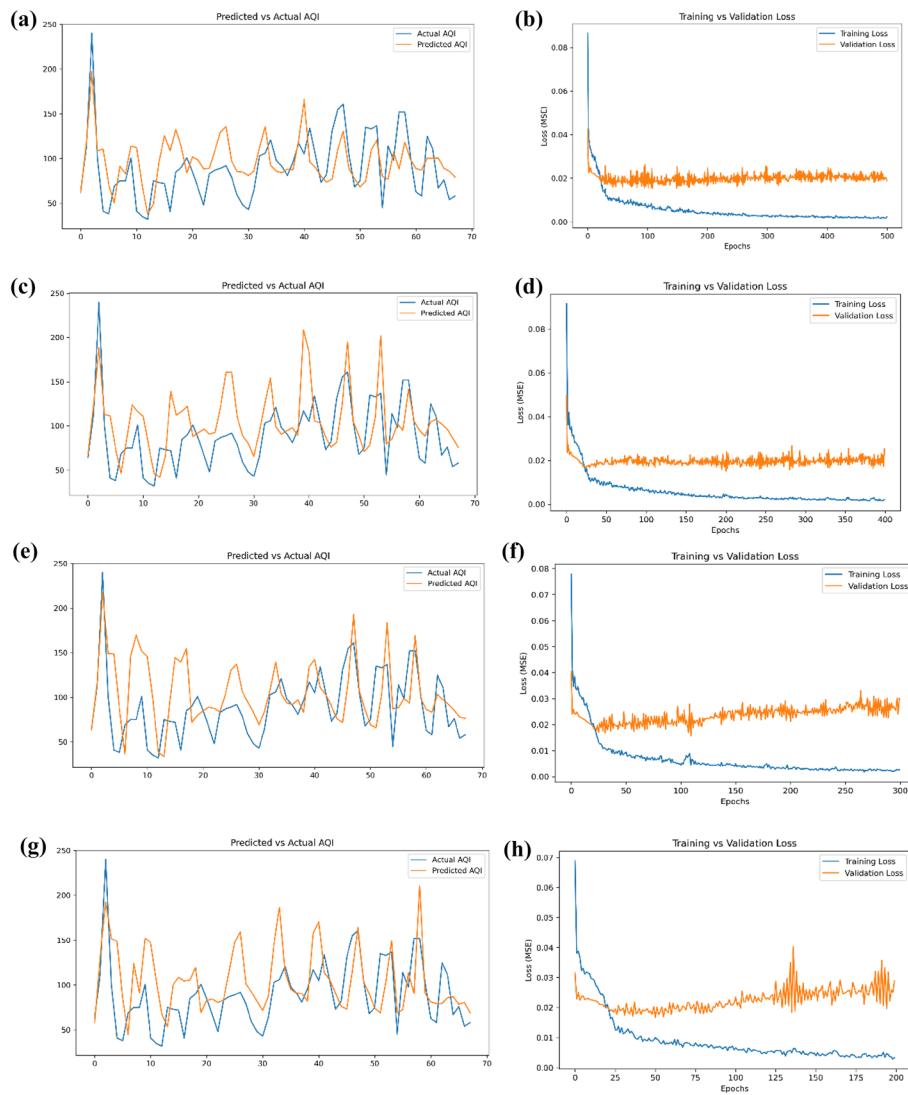


Fig. 25 **a** Actual and predicted graph of Adam optimizer at 500 epochs of Odisha **b** Training and validation loss graph of Adam optimizer at 500 epochs of Odisha **c** Actual and predicted graph of Adam optimizer at 400 epochs of Odisha **d** Training and validation loss graph of Adam optimizer at 400 epochs of Odisha **e** Actual and predicted graph of Adam optimizer at 300 epochs of Odisha **f** Training and validation loss graph of Adam optimizer at 300 epochs of Odisha **g** Actual and predicted graph of Adam optimizer at 200 epochs of Odisha **h** Training and validation loss graph of Adam optimizer at 200 epochs of Odisha

6 Results and discussion

With more precision and stability, the Adam optimizer outperformed RMSprop and SGD in this study on all key performance metrics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). Finally, the process ends after evaluating the model performance in AQI prediction. Adam achieved the lowest MAE of 8.38 at 200 epochs, followed by RMSprop with 8.63, while SGD produced significantly higher values. Table 3 also shows Adam's superiority over the other optimizers in reducing large and small prediction errors, with MSE and RMSE equal to 116.99 and 10.81, respectively, at 100 epochs. Table 5 MSE, RMSE, and MAE values of Adam, RMSprop, and SGD optimizers at 100, 200, and 300 epochs.

The above Table 5 illustrates the MSE, RMSE, and MAE values of RMSprop, and SGD optimizers at 100, 200, and 300 epochs. This result matches other research using deep

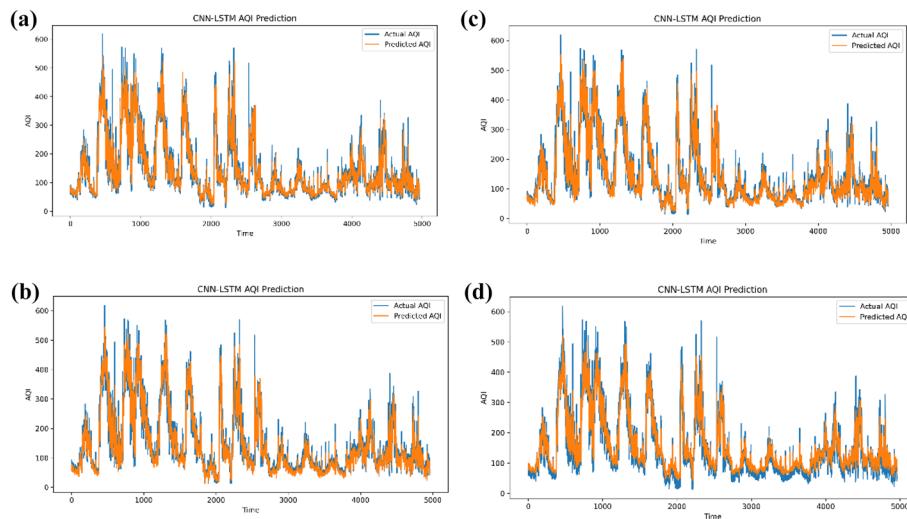


Fig. 26 **a** Actual and predicted graph of Adam optimizer at 20 epochs of city data of India **b** Actual and predicted graph of Adam optimizer at 40 epochs of city data of India **c** Actual and predicted graph of Adam optimizer at 60 epochs of city data of India **d** Actual and predicted graph of Adam optimizer at 80 epochs of city data of India.

Table 5 MSE, RMSE and MAE value of RMSprop and SGD optimizer at 100,200 and 300 epoch

Optimizers	Epoch	MSE	RMSE	MAE
RMS prop	100	155.81	12.48	9.66
	200	132.87	11.52	8.63
	300	153.20	12.377	9.26
SGD	100	1159	34.05	26.89
	200	148.6	12.19	9.39
	300	153.15	12.37	8.84

learning models to predict air quality. For example [64], used an Adam-optimized CNN-LSTM model to predict PM_{2.5} levels in Beijing. They achieved an RMSE of 12.4, which is much higher than the RMSE in this study and may indicate that the model was better tuned or that the input data was less variable. For air pollutant forecasting in Delhi, Krishan et al. [36] employed LSTM using RMSprop and obtained an MAE of 10.2, greater than the MAE obtainable with Adam for this study. This demonstrates how well Adam handles complex time-dependent relationships and variability in environmental datasets. On the other hand, SGD consistently had the worst performance, registering the highest values for every error measure, which is in line with its slower convergence and sensitivity to local minima in high-dimensional architectures like CNN-LSTM. Adam is the best optimizer for the CNN-LSTM model used in this study, offering the optimum trade-off between time needed for training, prediction accuracy, and stability based on these results and related literature.

Table 6 shows the performance of the model metrics (MSE, RMSE, MAE). In Anugul, the model had a minimum of RMSE equal to 10.81 and MAE equal to 8.38, with its stable performance throughout epochs. At Anugul, the model had a minimum RMSE of 10.81 and MAE of 8.38, with consistent performance across epochs. At the state level (Odisha), RMSE values converged around 35–40, with the best model performance at epoch 500 (RMSE = 34.53, MAE = 27.65). For India, the model uniformly generated RMSE values around 37–39 with comparatively low MAE values (24–26), showing its resilience across scales.

Table 6 Results tables having the MSE, RMSE, and MAE values for Anugul, Odisha and India data at different epochs for Adam optimizer

Data set	Epoch	MSE	RMSE	MAE
Anugul	100	116.99	10.81	8.53
	200	130.66	11.4	8.38
Odisha	300	131.30	11.45	8.85
	100	1408.53	37.53	31.69
	200	1565.78	39.57	31.40
	300	1630.64	40.38	30.99
	400	1373.65	37.06	29.87
	500	1192	34.53	27.65
India	1000	1534.35	39.17	32.67
	1500	1638	40.48	31.77
	2000	1385.59	37.22	29.42
	2500	1284.41	35.84	29.19
	3000	1287.49	35.88	27.92
	60	1425	37.75	23.81
India	80	1430.12	37.81	26.71
	500	1567.36	39.59	24.71

Conventional statistical models like ARIMA (1,1,1) and SARIMA were also tested for comparison. ARIMA models generally reported higher rates of errors with RMSE values ranging from 45 to 55 and MAE values of more than 30. In contrast, SARIMA performed marginally better but still lagged behind CNN-LSTM. This shows that the CNN-LSTM model learns the spatio-temporal relationships in AQI data better and performs consistently better than conventional time-series models, providing a more robust quantitative rationale for AQI prediction.

6.1 One-way ANOVA finding pertaining to CNN-LSTM architecture

ANOVA is a parametric test that depends on three distributional assumptions: (a) study group scores must be independent; (b) the distribution of each group's scores must be regular (normality); (c) the variances of group scores must be equal or constant [33]. The F-value in a one-way ANOVA informs us how much the group means vary compared to the group variability. Equation 1 is the formula for the F-value in one-way ANOVA.

$$F \text{ value} = BGV/WGV \quad (1)$$

Where, BGV (Between-Group Variance)= To what extent do the means of the dataset differ (Anugul vs. Odisha vs. India).

WGV (Within-Group Variance): To what extent the values within each dataset differ from its own dataset mean.

F-value: $Ratio = BGV \div WGV$

P-value: The probability of such differences happening.

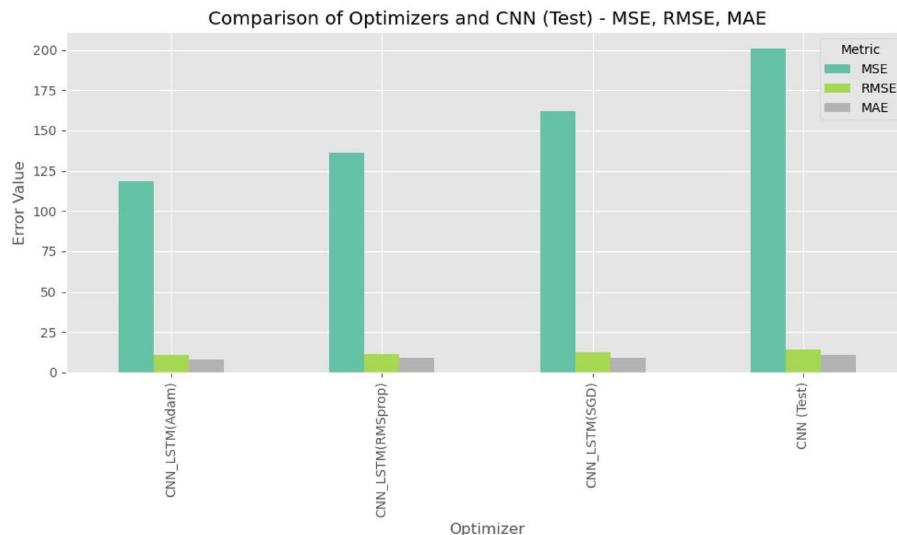
The difference between datasets (Anugul vs. Odisha vs. India) is much larger than within each dataset → Robust evidence that dataset means differ significantly.

F-value (almost 1) → Between-group and within-group variations are equal → little or no evidence for differences between datasets.

Table 7 below presents the outputs of the one-way ANOVA assessment conducted for three datasets (Anugul, Odisha, India) using the CNN-LSTM framework.

Table 7 Outputs of the one-way ANOVA

Metrics	F-Value	p-Value	Interpretation
MSE	260.54	1.23e-12	Significant ($p < 0.05$)
RMSE	2595.79	2.11e-23	Significant ($p < 0.05$)
MAE	770.43	1.74e-17	Significant ($p < 0.05$)

**Fig. 27** Comparison of optimizers

In Table 7 above, for all three metrics (MSE, RMSE, MAE), p-values are well below 0.05. Therefore, statistically significant differences exist between CNN-LSTM model performance for the datasets (Anugul, Odisha, India). Outputs of one-way ANOVA ($p < 0.05$ for all the error measures) validate that the predictive performance of the CNN-LSTM model substantially varies across the datasets (Anugul, Odisha, India). This means that while CNN-LSTM is a good model for AQI prediction, its performance would be largely place-based. Therefore, CNN-LSTM must be individually trained and tested for each location to give precise AQI forecasts.

The comparison graph indicates the performance of different model specifications, CNN-LSTM models trained with different optimizers (Adam, RMSprop, SGD), and a single CNN on three error metrics, Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) (see Fig. 27). The CNN-LSTM model trained using Adam has the minimum error values across all measures, indicating its highest predictive precision and generalizability. On the other hand, the same model with SGD and RMSprop training has moderately higher errors, and SGD has the worst performance of the three. The CNN model in isolation has the highest errors, indicating how powerful it is to merge sequential and convolutional learning in the CNN-LSTM model. Overall, the results emphasize the importance of selecting both optimizer and model structure, and that Adam-optimized CNN-LSTM is the optimal approach for AQI prediction. These comparisons show the capability of the CNN-LSTM model trained with the Adam optimizer in preserving spatial-temporal relationships and reducing wrong predictions. This model's much lower MAE value reveals that it can provide more accurate and reliable AQI level prediction, making it the top candidate

for deployment in real-world applications for monitoring pollution systems [30]. demonstrates that a low MAE means that model predictions are closer to the real value with rare material errors.

7 Conclusion

This article proposes a novel CNN-LSTM model to improve air quality prediction. It predicts air quality index levels based on Anugul City data from January 2019 to December 2023 to address some serious issues related to conventional prediction methods. The paper examines the Anugul area's air quality forecast and proposes ways to enhance it. By finding the temporal and spatial relationships between air pollution data, a hybrid CNN-LSTM model-based AQI predictive analysis shows a significant improvement in accuracy in prediction. The suggested model boosts the predictive accuracy and provides interpretability by incorporating Partial Dependence Plots (PDPs). The PDPs identify the separate impacts of major pollutants SO_2 , NO_2 , NH_3 , O_3 , PM_{10} , and $\text{PM}_{2.5}$ on AQI prediction. The PDP analysis highlights the significant role played by PM_{10} and $\text{PM}_{2.5}$ and determines that these are the key drivers of AQI fluctuations in the area. One of the significant contributions of this work is the comparison of optimization algorithms (Adam, RMSprop, and SGD) employed to train the CNN-LSTM model. Comparison parameters like Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) were employed for various training epochs (100, 200, 300) to determine model performance. Out of all the combinations, the Adam optimizer at 200 epochs yielded the minimum MAE (8.38), reflecting the maximum prediction accuracy and the best generalization ability for AQI prediction in the industrial area of Anugul. Conversely, the SGD optimizer was weak, especially at lower epochs, and reflects its inappropriateness for this use. Results verify the robustness and practical applicability of the Adam-optimized CNN-LSTM model as an effective and viable tool for real-time air pollution monitoring systems.

Although CNN-LSTM is an established architecture, the contribution of this work comes from its explainability part and its use in the industrially dominated area of Anugul. In our experiment, the Anugul dataset yielded the best predictive performance with a lowest RMSE of 10.81 and MAE of 8.38, showing that the model can capture localized spatio-temporal dynamics well. For the state level (Odisha), RMSE values were greater, ranging from 34 to 40, with the best result at epoch 500 ($\text{RMSE} = 34.53$, $\text{MAE} = 27.65$), and this implies that higher regional variability and heterogeneous pollution sources complicate prediction. For the national dataset (India), the model had comparatively stable RMSE values (37–39) with MAE of 24–26, consistent robustness, and indicated that spatial aggregation reduces fine-grained precision compared to local prediction. This comparative analysis shows that the CNN-LSTM works best in local settings (Anugul), is consistent at the state level (Odisha), and generalizes fairly well to country-level data, thus highlighting its flexibility as well as the necessity of incorporating explainability tools for effective decision-making.

The findings also offer valuable insights for policymakers and environmental agencies to come up with evidence-based, specific strategies for air pollution control, with PM_{10} and $\text{PM}_{2.5}$ emissions being the prime targets since the research found these to be the most critical determinants of air quality.

Particularly, air quality in industrial regions strongly influences the health impacts of neighbouring communities since sustained exposure to increased concentrations of pollutants, such as particulate matter, can result in severe respiratory and cardiovascular complications. Hence, incorporating predictive models such as the suggested CNN-LSTM into monitoring systems in industrial regions can significantly improve early warning capacities and reduce health risks for populations in neighbouring areas.

Acknowledgements

The authors are very thankful to the Siksha 'O' Anusandhan Deemed to be University for their encouragement and constant support.

Author contributions

E. M. Conceptualization, Methodology, Software, Validation, Investigation, Formal Analysis, Writing—original draft, Writing- review and editing. M. D. Conceptualization, Writing-review and editing. S.R. Conceptualization, Writing-review and editing, Software Validation. All authors reviewed the manuscript.

Funding

Open access funding provided by Siksha 'O' Anusandhan (Deemed To Be University). No funding available.

Data availability

Data available at this website : [<https://ospcboard.org/environmental-monitoring-data/>] (<https://ospcboard.org/environmental-monitoring-data>), [<https://cpcb.nic.in/National-Air-Quality-Index>] (<https://cpcb.nic.in/National-Air-Quality-Index>).

Declarations

Ethics approval and consent to participate

The authors are aware of the research ethics. All authors have read, understood, and complied as applicable with the statement on "Ethical responsibilities of Authors" as found in the Instructions for Authors and are aware that with minor exceptions, no changes can be made to authorship once the paper is submitted. not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 23 July 2025 / Accepted: 29 September 2025

Published online: 04 November 2025

References

- 1 Allgaier J, Mulansky L, Draelos RL, Pryss R. How does the model make predictions? A systematic literature review on the explainability power of machine learning in healthcare. *Artif Intell Med*. 2023;143:102616. <https://doi.org/10.1016/j.artmed.2023.102616>.
- 2 Ansari A, Quaff AR. Data-driven analysis and predictive modelling of hourly air quality index (AQI) using deep learning techniques: a case study of Azamgarh, India. *Theoret Appl Climatol*. 2025;156(1):74. <https://doi.org/10.1007/s00704-024-05304-y>.
- 3 Ansari N, Kumar P, Kumar R, Kumar P, Shamshad A, Hossain S, Sharma A, Singh Y, Kumari M, Mishra VN, Rukhsana, Javed A. Seasonal patterns of air pollution in delhi: interplay between meteorological conditions and emission sources. *Environ Geochem Health*. 2025;47(5):175. <https://doi.org/10.1007/s10653-025-02474-0>.
- 4 Barthwal A, Goel AK. Advancing air quality prediction models in urban india: a deep learning approach integrating DCNN and LSTM architectures for AQI time-series classification. *Model Earth Syst Environ*. 2024;10(2):2935–55. <https://doi.org/10.1007/s40808-023-01934-9>.
- 5 Bekkar A, Hssina B, Douzi S, Douzi K. Air-pollution prediction in smart city, deep learning approach. *J Big Data*. 2021;8(1):161. <https://doi.org/10.1186/s40537-021-00548-1>.
- 6 Bosson-Amedenu S, Baah EM, Ayiah-Mensah F, Addor JA. Advanced time series modelling to address seasonal variability and extreme rainfall in Ghana. *BMC Environ Sci*. 2025;2(1):8. <https://doi.org/10.1186/s44329-025-00024-8>.
- 7 Chang J, Hou J, Gong H, Sun Y. Optimization of CNN-LSTM air quality prediction based on the POA algorithm. *Sustainability*. 2025;17(12):5347. <https://doi.org/10.3390/su17125347>.
- 8 Dai Y, Wei J, Qin F. Recurrent neural network (RNN) and long short-term memory neural network (LSTM) based data-driven methods for identifying cohesive zone law parameters of nickel-modified carbon nanotube reinforced sintered nano-silver adhesives. *Mater Today Commun*. 2024;39:108991. <https://doi.org/10.1016/j.mtcomm.2024.108991>.
- 9 Dowlatabadi Y, Khajeh ZE, Mohammadi M, Sarkhosh M, Mohammad S, Moezzi M. Assessment of meteorological factors and air pollution impact on cardiovascular mortality using random forest analysis 2017 to 2020. *Sci Rep*. 2024;14(1):31468. <https://doi.org/10.1038/s41598-024-83185-x>.
- 10 Duan J, Gong Y, Luo J, Zhao Z. Air-quality prediction based on the ARIMA-CNN-LSTM combination model optimized by Dung beetle optimizer. *Sci Rep*. 2023;13(1):12127. <https://doi.org/10.1038/s41598-023-36620-4>.

- 11 Ekka S, Sahu SK, Dwivedi S, Khuman SN, Das S, Gaonkar O, Chakraborty P. Seasonality, atmospheric transport and inhalation risk assessment of polycyclic aromatic hydrocarbons in PM_{2.5} and PM₁₀ from industrial belts of Odisha, India. *Environ Geochem Health.* 2022;44(11):3991–4005. <https://doi.org/10.1007/s10653-021-01128-1>.
- 12 Elbeltagi A, Vishwakarma DK, Katipoğlu OM, Sushanth K, Heddam S, Singh BP, Shukla A, Gautam VK, Pande CB, Hussain S, Ghosh S, Dehghanianj H, Salem A. Air temperature Estimation and modeling using data driven techniques based on best subset regression model in Egypt. *Sci Rep.* 2025;15(1):2020. <https://doi.org/10.1038/s41598-025-06277-2>.
- 13 Gilik A, Ogrenci AS, Ozmen A. Air quality prediction using CNN + LSTM-based hybrid deep learning architecture. *Environ Sci Pollut Res.* 2022;29(8):11920–38. <https://doi.org/10.1007/s11356-021-16227-w>.
- 14 Godasaei SH, Ejohwomu OA, Zhong H, Booker D. Integrating experimental analysis and machine learning for enhancing energy efficiency and indoor air quality in educational buildings. *Build Environ.* 2025;276:112874. <https://doi.org/10.1016/j.buildenv.2025.112874>.
- 15 Gong Y, Wang Y, Xie Y, Peng X, Peng Y, Zhang W. Dynamic fusion LSTM-Transformer for prediction in energy harvesting from human motions. *Energy.* 2025;327:136192. <https://doi.org/10.1016/j.energy.2025.136192>.
- 16 Gu J, Yang B, Brauer M, Zhang KM. Enhancing the evaluation and interpretability of Data-Driven air quality models. *Atmos Environ.* 2021;246:118125. <https://doi.org/10.1016/j.atmosenv.2020.118125>.
- 17 Guo Q, He Z, Li S, Li X, Meng J, Hou Z, Liu J, Chen Y. Air pollution forecasting using artificial and wavelet neural networks with meteorological conditions. *Aerosol Air Qual Res.* 2020;20(6):1429–39. <https://doi.org/10.4209/aaqr.2020.03.0097>.
- 18 Guo Q, He Z, Wang Z. Change in air quality during 2014–2021 in Jinan City in China and its influencing factors. *Toxics.* 2023a;11(3):210. <https://doi.org/10.3390/toxics11030210>.
- 19 Guo Q, He Z, Wang Z. Predicting of daily PM_{2.5} concentration employing wavelet artificial neural networks based on meteorological elements in Shanghai, China. *Toxics.* 2023b;11(1):51. <https://doi.org/10.3390/toxics11010051>.
- 20 Guo Q, He Z, Wang Z. Prediction of hourly PM_{2.5} and PM₁₀ concentrations in Chongqing City in China based on artificial neural network. *Aerosol Air Qual Res.* 2023c;23(6):220448. <https://doi.org/10.4209/aaqr.220448>.
- 21 Guo Q, He Z, Wang Z. Monthly climate prediction using deep convolutional neural network and long short-term memory. *Sci Rep.* 2024a;14(1):17748. <https://doi.org/10.1038/s41598-024-68906-6>.
- 22 Guo Q, He Z, Wang Z. The characteristics of air quality changes in Hohhot City in China and their relationship with meteorological and Socio-economic factors. *Aerosol Air Qual Res.* 2024b;24(5):230274. <https://doi.org/10.4209/aaqr.230274>.
- 23 Guo Q, He Z, Wang Z, Qiao S, Zhu J, Chen J. A performance comparison study on climate prediction in Weifang City using different deep learning models. *Water.* 2024;16(19):2870. <https://doi.org/10.3390/w16192870>.
- 24 Guo Q, Wang Z, He Z, Li X, Meng J, Hou Z, Yang J. Changes in air quality from the COVID to the Post-COVID era in the Beijing-Tianjin-Tangshan region in China. *Aerosol Air Qual Res.* 2021;21(12):210270. <https://doi.org/10.4209/aaqr.210270>.
- 25 Gupta NS, Mohta Y, Heda K, Armaan R, Valarmathi B, Arulkumaran G. Prediction of air quality index using machine learning techniques: A comparative analysis. *J Environ Public Health.* 2023;2023:1–26. <https://doi.org/10.1155/2023/4916267>.
- 26 Hasan MM. Understanding model predictions: A comparative analysis of SHAP and LIME on various ML algorithms. *J Sci Technological Res.* 2024;5(1):17–26. [https://doi.org/10.59738/jstr.v5i1.23\(17-26\).eaqr5800](https://doi.org/10.59738/jstr.v5i1.23(17-26).eaqr5800).
- 27 Hasan P, Khan TD, Alam I, Haque ME. Dengue in tomorrow: predictive insights from ARIMA and SARIMA models in Bangladesh: a time series analysis. *Health Sci Rep.* 2024;7(12). <https://doi.org/10.1002/hsr.2.7026>.
- 28 He Z, Guo Q. Comparative analysis of multiple deep learning models for forecasting monthly ambient PM_{2.5} concentrations: A case study in Dezhou City, China. *Atmosphere.* 2024;15(12):1432. <https://doi.org/10.3390/atmos15121432>.
- 29 He Z, Guo Q, Wang Z, Li X. Prediction of monthly PM_{2.5} concentration in Liaocheng in China employing artificial neural network. *Atmosphere.* 2022;13(8):1221. <https://doi.org/10.3390/atmos13081221>.
- 30 Hodson TO. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geosci Model Dev.* 2022;15(14):5481–7. <https://doi.org/10.5194/gmd-15-5481-2022>.
- 31 Hu Y, Ding Y, Jiang W. Geographically aware air quality prediction through CNN-LSTM-KAN hybrid modeling with Climatic and topographic differentiation. *Atmosphere.* 2025;16(5):513. <https://doi.org/10.3390/atmos16050513>.
- 32 Iqbal A, Mukherjee N. A systematic review and comparative study of machine learning techniques for air quality prediction. *Water Air Soil Pollut.* 2025;236(12):789. <https://doi.org/10.1007/s11207-025-08410-3>.
- 33 Juarros-Basterretxea J, Alonso-Diego G, Postigo Á, Montes-Alvarez P, Menéndez-Aller Á, García-Cueto E. Post-hoc tests in one-way ANOVA: the case for normal distribution. *Methodology.* 2024;20(2):84–99. <https://doi.org/10.5964/meth.11721>.
- 34 Khashei M, Bijari M, Hejazi SR. Combining seasonal ARIMA models with computational intelligence techniques for time series forecasting. *Soft Comput.* 2012;16(6):1091–105. <https://doi.org/10.1007/s00500-012-0805-9>.
- 35 Kim Y, Kim Y. Explainable heat-related mortality with random forest and SHapley additive explanations (SHAP) models. *Sustainable Cities Soc.* 2022;79:103677. <https://doi.org/10.1016/j.scs.2022.103677>.
- 36 Krishan M, Jha S, Das J, Singh A, Goyal MK, Sekar C. Air quality modelling using long short-term memory (LSTM) over NCT-Delhi, India. *Air Qual Atmos Health.* 2019;12(8):899–908. <https://doi.org/10.1007/s11869-019-00696-7>.
- 37 Li T, Hua M, Wu X. A hybrid CNN-LSTM model for forecasting particulate matter (PM_{2.5}). *IEEE Access.* 2020;8:26933–40. <https://doi.org/10.1109/ACCESS.2020.2971348>.
- 38 Lim B, Zohren S. Time-series forecasting with deep learning: a survey. *Philosophical Trans Royal Soc A: Math Phys Eng Sci.* 2021;379(2194):20200209. <https://doi.org/10.1098/rsta.2020.0209>.
- 39 Llamelo C, Medina R, Fajardo A. Performance of enhanced CNN-LSTM prediction model for vehicle-mounted air quality monitoring. 2024 15th International Conference on Information and Communication Technology Convergence (ICTC), 66–71. 2024. <https://doi.org/10.1109/ICTC62082.2024.10826663>.
- 40 M P. Impact of environmental pollution on respiratory system of human and animals in Angul and Talcher industrial Areas, Odisha, India: A case study. *Int J Zool Anim Biology.* 2019;2(4):1–15. <https://doi.org/10.23880/IJZAB-16000162>.
- 41 McDonnell K, Murphy F, Sheehan B, Masello L, Castignani G. Deep learning in insurance: accuracy and model interpretability using TabNet. *Expert Syst Appl.* 2023;217:119543. <https://doi.org/10.1016/j.eswa.2023.119543>.
- 42 Mills TC. Modelling stationary time series: the ARMA Approach. In *Time series econometrics.* Palgrave Macmillan UK;2015. pp. 5–40. https://doi.org/10.1057/9781137525338_2.
- 43 Mishra A, Gupta Y. Comparative analysis of air quality index prediction using deep learning algorithms. *Spat Inform Res.* 2024;32(1):63–72. <https://doi.org/10.1007/s41324-023-00541-1>.
- 44 Molnar C. *Interpretable machine learning.* Lulu. com; 2020.

- 45 Mubeen M, He S, Rahman MS, Wang L, Zhang X, Ahmed B, He Z, Han Y. Smart prediction and optimization of air quality index with artificial intelligence. *J Environ Sci.* 2025. <https://doi.org/10.1016/j.jes.2025.02.041>.
- 46 Muhammad LJ, Haruna AA, Sharif US, Mohammed MB. CNN-LSTM deep learning based forecasting model for COVID-19 infection cases in Nigeria, South Africa and Botswana. *Health Technol.* 2022;12(6):1259–76. <https://doi.org/10.1007/s1253-022-00711-5>.
- 47 Nguyen AT, Pham DH, Oo BL, Ahn Y, Lim BTH. Predicting air quality index using attention hybrid deep learning and quantum-inspired particle swarm optimization. *J Big Data.* 2024;11(1):71. <https://doi.org/10.1186/s40537-024-00926-5>.
- 48 Organization WH et al. WHO global air quality guidelines: particulate matter (PM_{2.5} and PM₁₀), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization. 2021.
- 49 Özüpak Y, Alpsalaz F, Aslan E. Air quality forecasting using machine learning: comparative analysis and ensemble strategies for enhanced prediction. *Water Air Soil Pollut.* 2025;236(7):464. <https://doi.org/10.1007/s11270-025-08122-8>.
- 50 Raj S, Smith J, Hayes E. Hybrid graph convolutional LSTM model for spatio-temporal air quality transfer learning. *Air Qual Atmos Health.* 2025;18(5):1425–45. <https://doi.org/10.1007/s11869-025-01713-8>.
- 51 Roy T, Das P, Jagirdar R, Shhabat M, Abdullah MS, Kashem A, Rahman R. Prediction of mechanical properties of eco-friendly concrete using machine learning algorithms and partial dependence plot analysis. *Smart Constr Sustainable Cities.* 2025;3(1):2. <https://doi.org/10.1007/s44268-025-00048-8>.
- 52 Salim M, Djunaidy A. Development of a CNN-LSTM approach with images as Time-Series data representation for predicting gold prices. *Procedia Comput Sci.* 2024;234:333–40. <https://doi.org/10.1016/j.procs.2024.03.007>.
- 53 Sarkar N, Gupta R, Keserwani PK, Govil MC. Air quality index prediction using an effective hybrid deep learning model. *Environ Pollut.* 2022;315:120404. <https://doi.org/10.1016/j.envpol.2022.120404>.
- 54 Saxena V. Water Quality, air Pollution, and climate change: investigating the environmental impacts of industrialization and urbanization. *Water Air Soil Pollut.* 2025;236(2):73. <https://doi.org/10.1007/s11270-024-07702-4>.
- 55 Su Y, Li J, Liu L, Guo X, Huang L, Hu M. Application of CNN-LSTM algorithm for PM2.5 concentration forecasting in the Beijing-Tianjin-Hebei metropolitan area. *Atmosphere.* 2023;14(9):1392. <https://doi.org/10.3390/atmos14091392>.
- 56 Sun F, Hao W, Zou A, Shen Q. A survey on spatio-temporal series prediction with deep learning: taxonomy, applications, and future directions. *Neural Comput Appl.* 2024;36(17):9919–43. <https://doi.org/10.1007/s00521-024-09659-1>.
- 57 Szostek K, Mazur D, Draflus G, Kusznier J. Analysis of the effectiveness of ARIMA, SARIMA, and SVR models in time series forecasting: A case study of wind farm energy production. *Energies.* 2024;17(19):4803. <https://doi.org/10.3390/en17194803>.
- 58 Van den Broeck G, Lykov A, Schleich M, Suciu D. On the tractability of SHAP explanations. *J Artif Intell Res.* 2022;74:851–86. <https://doi.org/10.1613/jair.1.13283>.
- 59 Wang H, Liang Q, Hancock JT, Khoshgoftaar TM. Feature selection strategies: a comparative analysis of SHAP-value and importance-based methods. *J Big Data.* 2024;11(1):44. <https://doi.org/10.1186/s40537-024-00905-w>.
- 60 Wang J, Li X, Jin L, Li J, Sun Q, Wang H. An air quality index prediction model based on CNN-ILSTM. *Sci Rep.* 2022;12(1):8373. <https://doi.org/10.1038/s41598-022-12355-6>.
- 61 Wu Q, Lin H. A novel optimal-hybrid model for daily air quality index prediction considering air pollutant factors. *Sci Total Environ.* 2019;683:808–21. <https://doi.org/10.1016/j.scitotenv.2019.05.288>.
- 62 Yadav R, Sahu LK, Beig G, Tripathi N, Jaaffrey SNA. Ambient particulate matter and carbon monoxide at an urban site of India: influence of anthropogenic emissions and dust storms. *Environ Pollut.* 2017;225:291–303. <https://doi.org/10.1016/j.envpol.2017.01.038>.
- 63 Zhang J, Li S. Air quality index forecast in Beijing based on CNN-LSTM multi-model. *Chemosphere.* 2022;308:136180. <https://doi.org/10.1016/j.chemosphere.2022.136180>.
- 64 Zhang Z, Zeng Y, Yan K. A hybrid deep learning technology for PM_{2.5} air quality forecasting. *Environ Sci Pollut Res.* 2021;28(29):39409–22. <https://doi.org/10.1007/s11356-021-12657-8>.
- 65 Zhao Q, Liu J, Yang X, Qi H, Lian J. Spatiotemporal <math altimg=si5.svg display=inline id=d1e371> mrow > <mi>P</mi> M <mrow > <mtext>2.5</mtext> forecasting via dynamic geographical graph neural network. *Environ Model Softw.* 2025;186:106351. <https://doi.org/10.1016/j.envsoft.2025.106351>.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.