International Journal of Electrical and Computer Engineering (IJECE)
Vol. 16, No. 1, February 2026, pp. 333~341
ISSN: 2088-8708, DOI: 10.11591/ijece.v16i1.pp333-341

Optimizing hourly air quality index forecasting: a particle swarm optimization-enhanced hybrid approach combining convolutional and recurrent neural networks

Darakhshan Khan[1], Archana B. Patankar[1], Jyotika Kakar[2]
[1]Department of Computer Engineering, Thadomal Shahani Engineering College, University of Mumbai, Mumbai, India
[2]Department of Computer Science and Engineering, University of California, San Diego, United States

Article Info

Article history:

Received Aug 19, 2024
Revised Oct 16, 2025
Accepted Nov 23, 2025

Keywords:

Air quality index
Convolution
Deep learning
Long short-term memory
Particle swarm optimization

ABSTRACT

Air pollution is still a serious worldwide issue, and accurate air quality index (AQI) prediction is needed. This paper proposes a hybrid deep learning model integrating 1D convolutional neural networks (Conv1D) and long short-term memory (LSTM) networks, optimized with particle swarm optimization (PSO) to enhance AQI forecasting. The model was evaluated at six urban areas: Bandra, Thane, Mazgaon, Kurla, Nerul, and Malad, and compared with a single LSTM network. PSO adjusted hyperparameters like hidden units, batch size, epochs, and learning rate was used to improve predictive accuracy. The Conv1D+LSTM hybrid model drastically decreased RMSE by 49.19% (Bandra), 33.97% (Thane), 5.24% (Mazgaon), 20.52% (Kurla), 35.85% (Nerul), and 27.54% (Malad), and R^2 Score improvements up to 751.2%. Training logs indicated smoother convergence with loss decrease at faster rates compared to LSTM, showing better learning efficiency and generalization. By combining spatial and temporal feature extraction with automated hyperparameter tuning, this model captures sophisticated pollution patterns which increases the reliability of AQI prediction. Enhancements in the future can be adding regularization methods and more feature inputs to improve the accuracy.

Corresponding Author:

Darakhshan Khan
Department of Computer Engineering, Thadomal Shahani Engineering College, University of Mumbai
Mumbai, India
Email: darakshan.khan@thadomal.org

1. INTRODUCTION

Air pollution remains one of the most urgent global challenges, contributing to approximately seven million premature deaths each year [1], [2]. Effective monitoring and prediction of air quality are essential to mitigating its harmful impacts. The air quality index (AQI) serves as a standardized tool to convey air quality levels, empowering the public and policymakers to make informed decisions [3], [4]. However, despite the widespread deployment of air monitoring stations and the availability of extensive datasets, accurately forecasting AQI poses a significant challenge. This complexity arises from the intricate temporal and spatial patterns embedded within air quality data.

A wide range of articles have applied deep learning methods to enhance AQI prediction. The use of sequence-to-sequence temporal models, including recurrent neural networks (RNNs), long short-term memory (LSTM) networks, and gated recurrent units (GRUs), have been proven to be effective in time-series forecasting because of their capacity to model temporal dependencies [5], [6]. For multivariate AQI forecasting, different deep learning architectures have shown robustness in understanding the variation in time-series data. Multiple variables, such as, pollutant level, weather factors and spatial factors related to geographical location can be used for forecasting [7]. For instance, hybridized techniques, convolutional neural network-LSTM (CNN-LSTM) [8], [9] models have been applied in spatial temporal clustering to enhance the forecast of AQI at variety of places and time horizons. Wuhan is one case, where a fuzzy entropy based ensemble LSTM model combined with decomposition reconstruction techniques significantly enhances daily prediction correctness for AQI [10]. Similarly, Roy *et al.* [11] have depicted that the hybridization of LSTM models outclasses isolated models for the AQI prediction within Kattankulathur and Kolkata. In Sichuan, China, architectures which are using LSTM cells have surpassed traditional algorithms like back propagation (BP) and GRU networks [12].

Other revolutionary composite techniques have emerged in recent studies. Particularly, Prophet-LSTM models have demonstrated excellent accuracy for extreme AQI values forecast in Nanjing [13]. Wavelet transform and advanced transformer networks are shown to effectively record time-frequency domain parameters for AQI modeling in Guilin [14]. Random connectivity-enabled LSTM models further enhanced prediction reliability while reducing the cost of computation and surpassing standard approaches like support vector regression (SVR), autoregressive integrated moving average (ARIMA), and feedforward neural network (FFNN) [15]. Apart from the aforementioned methods, Gupta *et al.* [16] dealt with the issue of data imbalance by employing the synthetic minority oversampling technique (SMOTE) along with regression techniques, like CatBoost and SVR, for AQI forecasting in Indian cities. Those studies provided better multivariate feature handling. A review in depth [17], [18] have raised the awareness of ensemble approaches like stacking, bagging, and boosting simultaneously alongside more recent deep learning paradigms, along with their prospects and pitfalls.

Random forest was the preferred method for spatial modelling of $PM_{2.5}$, and deep learning, LSTM, XGBoost, and ensemble models followed, with advancement in machine learning create a window for improvement, inter-comparison studies, and new applications in unknown areas [19]. Optimization algorithms such as particle swarm optimization (PSO) have been crucial in improving model performance. Huang *et al.* [20] used PSO to optimize the hyperparameters of the LSTM for predicting water quality in China's South-North water transfer project, where substantial accuracy enhancement was observed. Similarly, a CNN-Bi-LSTM hybrid model optimized with PSO revealed lesser error values such as mean absolute error (MAE), mean squared error (MSE), and root mean squared error (RMSE) for the wheat yield forecast [21]. These advancements go to underscore the efficiency of PSO and other nature-inspired algorithms to improve the accuracy and effectiveness of deep learning-based forecasting models.

In spite of major breakthroughs in deep learning for AQI prediction, most present models lack flexibility across varied urban environments because feature extraction is not efficient and hyperparameter tuning process is not optimal. Most methods are based on either CNN for spatial feature extraction or LSTMs for sequential modeling and don't take advantage of the dual strengths of both architectures. Moreover, hyperparameter optimization of deep learning models is usually done manually, thus rendering process is cumbersome and reduces predictive precision. Although there has been significant improvement in AQI prediction, there are still several challenges, including guaranteeing model stability across different geographic locations, which impacts generalizability and real-world usability.

In addition, hyperparameter optimization, a key step towards improving model performance, is largely uninvestigated in most of the previous work. Most current methods are focused on temporal or spatial aspects but do not combine both with great effectiveness and hence result in less accurate partial forecasts. Filling these blanks, this research proposes a hybrid Conv1D+LSTM based on PSO that not only enhances feature presentation but also, through automatic optimization of hyperparameters, greatly promotes forecasting accuracy. Through the model's validation across six different urban sites, the research proves its better predictive capacity than individual deep learning models, supporting the requirement for more generalized and optimized solutions to address the multifaceted challenges of AQI prediction under various environmental conditions.

Effective AQI forecasting involves modeling both spatial and temporal dependencies in air pollution data, which is not often handled well by conventional models. To overcome such shortcomings, this research introduces a hybrid deep learning framework involving 1D convolutional neural networks (Conv1D) and LSTM networks, tuned with PSO. The Conv1D layer learns spatial features from air pollution data and the LSTM layer learns temporal relationships so that the model can identify long-term changes in air quality trends. PSO again improves the performance of the model by fine-tuning the hyperparameters including the number of neurons, batch size, learning rate, and epochs so that the network is optimized for every dataset. The robustness and generalizability of the model are ensured by validation over six different locations in Mumbai, which are Bandra, Thane, Mazgaon, Kurla, Nerul, and Malad; demonstrating its capability to give an accurate and reliable AQI prediction over different environmental conditions.

The following sections provide a detailed outline of the methodology and validate the relevance of the proposed approach. Section 2 describes the experimental setup, including data collection and preprocessing procedures, the design of the hybrid Conv1D-LSTM model architecture, and the evaluation metrics used to measure its performance. Section 3 offers a comprehensive analysis of the results, demonstrating the hybrid model's superiority through comparisons with existing methods across various metrics and geographic locations. Lastly, section 4 concludes the study by summarizing the key findings and presenting potential future directions to further enhance the accuracy and applicability of AQI forecasting.

2. METHOD

This section further explains each step that was used in this research, as shown in Figure 1, which includes data acquisition, data handling followed by use of PSO algorithm for finding optimal hyper parameter, next step is data transformation and division, and lastly, model training and testing.

2.1. Data collection

This dataset includes AQI readings for six sites, namely, Bandra, Thane, Mazgaon, Kurla, Nerul and Malad, which are located in the Indian city of Mumbai. The data has varied characteristics, ranging from pollutant concentrations such as $PM_{2.5}$, PM_{10}, NOx, SO_2, CO, Ozone; climatic characteristics like temperature, dew_point, temperature_min, temperature_max, and diurnal temperature, pressure, humidity, wind speed and wind degrees. The hourly frequency data of pollutants were taken from the CPCB [22] website, and meteorological parameters for different sites were retrieved using the OpenWeatherMap [23] API. Table 1, shows summary of data collected for six different sites with total number records per site where frequency of observations is hourly, that is, there are 24 values recorded in a day.

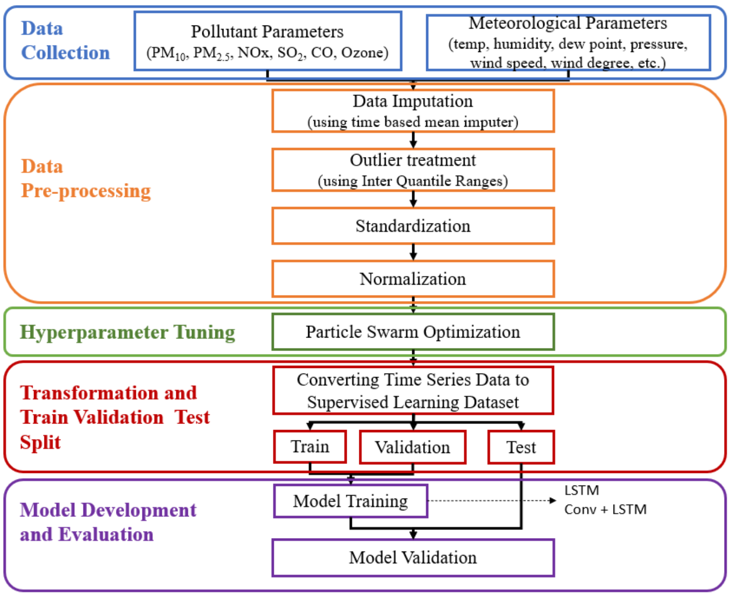


Figure 1. Overview of research methodology

Table 1. Detailed record count of data for six sites

S. No	Site	Number of Records	Time Period
1	Bandra	63529	01.01.2017 – 31.03.2024
2	Thane	63529	01.01.2017 – 31.03.2024
3	Mazgaon	29929	01.11.2020 – 31.03.2024
4	Kurla	42251	01.06.2019 – 31.03.2024
5	Nerul	29929	01.11.2020 – 31.03.2024
6	Malad	29929	01.11.2020 – 31.03.2024

2.2. Data pre-processing

Various data processing stepa are performed to assure data quality and consistency. Missing value imputation is conducted by filling missing values with average of data value from the same day using the previous or following years, as applicable, and if missing values still exist, linear interpolation with forward fill is used. Then, outliers are identified using the interquartile ranges [24]. Data is passed to

"StandardScaler" and "MinMaxScaler" module of Scikit-Learn after outlier treatment to ensure that all of the features have the same scale and to limit them to a particular range of 0 to 1.

2.3. Particle swarm optimizer for hyperparameter tuning

Inspired by the social behavior of creatures like flocking birds, particle swarm optimization (PSO) is a population-based evolutionary method. This operates by dispersing a large number of interactive particles throughout the search area. Every particle has the potential to solve the optimization problem. As the particles travel, arguments are updated according to previous errors and those in its neighborhood. Tracing and following particles in the population is a method that is necessary for the swarm to converge towards its optimal solution [20], [21], [25].

The following five or six hyperparameters as shown in Table 2, were determined depending on model architecture using the PSO algorithm. For each of these hyperparameters, a set of discrete values were used and the PSO algorithm has to find the best combination that minimizes loss. This optimizer was designed to run for ten particles, with each of them performing five iterations.

Table 2. Hyperparameter values used for optimization search

S. No	Hyperparameters	Values
1	Number of Units in Hidden Layer 1 (**N1**)	[32, 64, 128]
2	Number of Units in Hidden Layer 2 (**N2**)	[16, 32, 64]
3	Number of Units in Hidden Layer 3 (**N3**)	[16, 32, 64]
4	Epochs	[10, 20, 30, 40, 50]
5	Batch size	[128, 256, 512]
6	Learning rate	[0.001, 0.01]

2.4. Data transformation and splitting

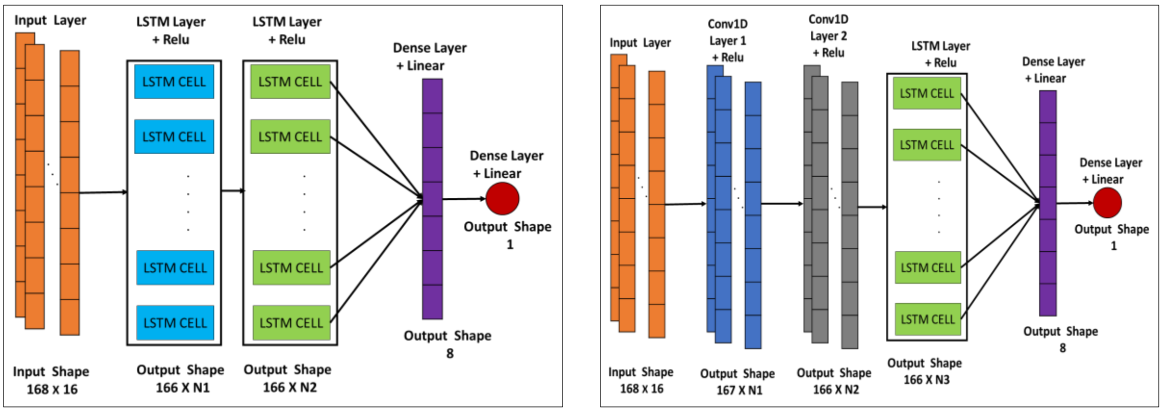
The normalized time series data is converted into supervised learning data set, which is then split into three distinct sets in a 70:20:10 basis: training, validation, and testing. The training set will be 70%, enabling it to learn the patterns and relationships within the dataset. It is during model development that the validation set, which is 20% of the data, is used to tune the model weights, so that the model does not overfit. Lastly, the remaining 10% goes to the testing set that is, used for model evaluation.

2.5. Model architectures

Two deep learning architectures implemented are LSTM and Conv1D+LSTM. One of the forms of RNN architecture, especially for the processing of sequential data, is the long short-term memory, or LSTM. LSTMs could decide whether information should be written, deleted, or read through a gating system within every cell of this network architecture. Input gate, forget gate, and output gate are three kinds of gates that are used in an LSTM cell [10], [26], [27]. The input gate states what information shall be added to this cell state and the output gate determines what information the LSTM should send out at that phase [2]. The Conv1D is a special type of convolution layer designed to handle one dimensional data such time series data. These Conv1D can extract local spatial features while processing input data sequences. In order to capture temporal dependencies in the retrieved features, the LSTM layer is applied after the Conv1D layer [28], [29]. The architecture diagram of the both networks used in this research is depicted in Figure 2. Figure 2(a) is the detailed LSTM architecture which consist of two stacked LSTM layer each of them has dropout regularization applied, and two dense layers towards the end. The input features are fed into stacked LSTM layers, shown in blue and green, to capture the temporal dependencies of the data. The output from the LSTM layers passes through fully connected dense layers, represented as purple blocks, to produce the final forecast.

Figure 2(b) is the architecture diagram of the proposed hybrid Conv1D+LSTM network. The model consists of two convolution layers, one LSTM layer, and two dense layers towards the end. The input features are fed into stacked 1D convolution layers, shown in blue and grey, to capture the spatial dependencies of the data and perform smoothing operation on data. The output from the Conv1D layers is passed to LSTM layer, shown in green, for understanding temporal correlation in data. Lastly, output from LSTM is fed to fully connected dense layers, represented as purple blocks, to produce the final forecast.

The optimum hyperparameter that the PSO algorithm searched for both networks and sites is shown in Table 3. The parameters are fine-tuned by PSO to improve model accuracy and efficiency in making accurate forecast. Values in Table 3 implies that LSTM has greater variability in hyperparameters (N1, batch size, epochs), which is more location-specific and requires greater fine-tuning. Whereas Conv1D+LSTM has more invariant hyperparameters between locations, meaning that the model structure is more generalized. For training, Conv1D+LSTM requires more epochs and larger batch sizes, which generally means higher computational cost, while LSTM would be trained faster with smaller sizes of a batch and epochs. Overall, the models use similar learning rate but LSTM appears to have stronger location-specific fine-tuning. Conv1D+LSTM remains relatively consistent along locations, indicating that LSTM is more sensitive to local patterns and requires finer adjustment of hyperparameters, and Conv1D+LSTM has provided a way with a little more robust and consistent framework.



(a) (b)

Figure 2. Architecture of (a) LSTM and proposed hybrid (b) Conv1D+LSTM network

Table 3. Optimized hyperparameter values identified by PSO algorithm

Models	Hyperparameters	Bandra	Thane	Mazgaon	Kurla	Nerul	Malad
LSTM	N1	64	128	32	64	128	32
	N2	16	16	16	16	16	16
	Epochs	30	20	40	20	40	20
	Batch Size	128	256	512	256	512	128
	Learning Rate	0.01	0.01	0.01	0.001	0.001	0.001
Conv1D+LSTM	N1	64	32	64	64	64	64
	N2	16	16	16	32	16	16
	N3	64	16	64	32	64	32
	Epochs	50	30	20	30	20	50
	Batch Size	512	256	256	128	128	512
	Learning Rate	0.001	0.01	0.001	0.01	0.01	0.01

2.6. Evaluation measures

The accuracy of the models is assessed using the mean absolute percentage error (MAPE), root mean square error (RMSE) [30] and coefficient of determination (R^2 Score) [6]. Equations (1) to (3) shows how these metrics are calculated using actual and predicted values, as follows,

$$MAPE = \frac{1}{N}\sum_{i=1}^{n}\left|\frac{(y_i - \hat{y})}{y_i}\right| \quad (1)$$

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

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y})^2}{\sum_{i=1}^{n}(y_i - \underline{y})^2} \quad (3)$$

Where N is the number of samples of the test set, y_i and $\hat{y}$ are the actual and predicted values at time instance i, respectively, and $\underline{y}$ is the mean of the test set sample values.

2.7. Simulation setup

All simulations for this research were carried out on Google Colaboratory, which is a Jupyter notebook-based environment on the cloud with GPU support. The models were developed and executed using Python 3.9 and a collection of open-source libraries such as TensorFlow 2.10, Keras 2.10, Scikit-learn 1.1.3, Pandas 1.3.5, NumPy 1.21.6, Matplotlib 3.5.1, and pyswarm 0.6 which is a library for PSO. Also, Colab's GPU runtime provided a Tesla T4 GPU with 12 GB RAM which improved the speed of the model training. In addition, during the simulation, a fixed random seed was used to ensure result reproducibility.

The models were trained with the Adam optimizer and the objective was to minimize mean squared logarithmic error (MSLE) loss. Convolution and LSTM layers used ReLU, and the output dense layers used a linear activation function. The simulation pipeline encompasses the entire workflow for data preprocessing, hyperparameter optimization using PSO, model training, and performance evaluation for each of the six locations, which are Bandra, Thane, Mazgaon, Kurla, Nerul, and Malad.

3. RESULT AND DISCUSSION

It can be observed from Table 4 that the Conv1D+LSTM model performed better than the stand-alone model with LSTM for all the sites. In particular, for Bandra and Thane, proposed hybrid architecture has outperform on all metrics. Even for remaining four stations results achieved by Conv1D+LSTM model have modest improvement over standalone LSTM architecture. For LSTM model and Conv1D+LSTM models, the actual and predicted values of AQI for three sites, namely, Bandra, Thane and Kurla are displayed in Figure 3, similar graphs were observed for other three sites. Actual values are shown in blue color in the graph, while forecasted values are shown in green color. Figures 3(a) and 3(b) are plots for Bandra site clearly indicates LSTM is good for capturing trend in AQI levels whereas hybrid model was robust in identifying sudden spikes and dips in the AQI levels. Similar observation can be seen for Thane site in Figures 3(c) and 3(d), where LSTM model was able to understand macro-trends and hybrid model could identify micro-pattern or localized trends. Likewise, Figures 3(e) and 3(f) shows results achieved for Kural site with exactly same observations. The proposed hybrid network clearly learned the sharp peaks in the dataset. That shows the hybrid model is much better in learning complex patterns and temporal dependencies underlying the data. The Conv1D component likely helps extract spatial features from the data, while the LSTM component does a good job in modeling temporal dependencies. The combination of these two can enable the model to exploit spatial and temporal information. The predicted values are very close to the ground-truth values, especially during sharp increases and decreases of AQI concentrations using proposed hybrid Conv1D+LSTM model than standalone LSTM model.

Figure 4 shows training history of LSTM and proposed model for three locations, in particular, Bandra, Thane, and Kurla sites. In these plots, blue line indicates RMSE values on training set after every epoch and orange line indicates RMSE values on validation set. On an average, training RMSE for LSTM models smoothens out after 13-14 epochs. Figures 4(a) to 4(c) are training logs for LSTM models where training RMSE curves in smooth indicating proper learning but validation RMSE curves are showing random fluctuations, which signifies models' inability in generalization as it is sensitive to noise and potential risk of overfitting. Similarly, Figures 4(d) to 4(f), represents training logs of Conv1D+LSTM model for three sites, Bandra, Thane and Kurla, similar results were seen for other three sites. After closely observing the graphs, it is evident that, there is smooth decrease in RMSE values for training data except for Bandra site where there is sudden spike possibly due improper hyper parameter settings. In contrast to LSTM models, hybrid models' validation RMSE has shown less irregularity except for Kurla. Overall, proposed hybrid Conv1D+LSTM's learning logs when compared with learning logs of LSTM model, it shows that there is smooth decrease in RMSE value for both training and validation set, exhibiting better generalization capabilities and they are more robust, but still, even for Conv1D+LSTM models training can be improved further by adjusting learning rates or using different regularization mechanisms.

The originality of this study is the combination of PSO with a Conv1D+LSTM hybrid model, which is not often utilized in AQI prediction. Contrary to other deep learning methods that need labor-intensive manual tuning, our PSO-based optimization procedure automates hyperparameter tuning, guaranteeing the best performance of the model with minimal human intervention. The hybrid framework also supports more accurate and generalizable AQI forecasts in different geographical locations, as seen in our experimental results.

Table 4. Performance measures, RMSE, R^2 Score and MAPE values for all LSTM and Conv1D+LSTM models for all sites

S. No	Site	RMSE		R^2 Score		MAPE	
		LSTM	Conv1D+LSTM	LSTM	Conv1D+LSTM	LSTM	Conv1D+LSTM
1	Bandra	43.1274	**21.9119**	-1.0021	**0.7620**	0.2981	**0.1163**
2	Thane	43.1458	**28.4916**	0.3633	**0.8587**	0.2161	**0.1041**
3	Mazgaon	23.0443	**21.8389**	0.2067	**0.5572**	0.1044	**0.0999**
4	Kurla	36.2694	**28.8264**	0.1297	**0.6639**	0.1847	**0.1382**
5	Nerul	63.3484	**40.6324**	-2.4412	**0.4224**	0.3706	**0.1735**
6	Malad	37.4798	**27.1554**	-0.1082	**0.7049**	0.2069	**0.1423**

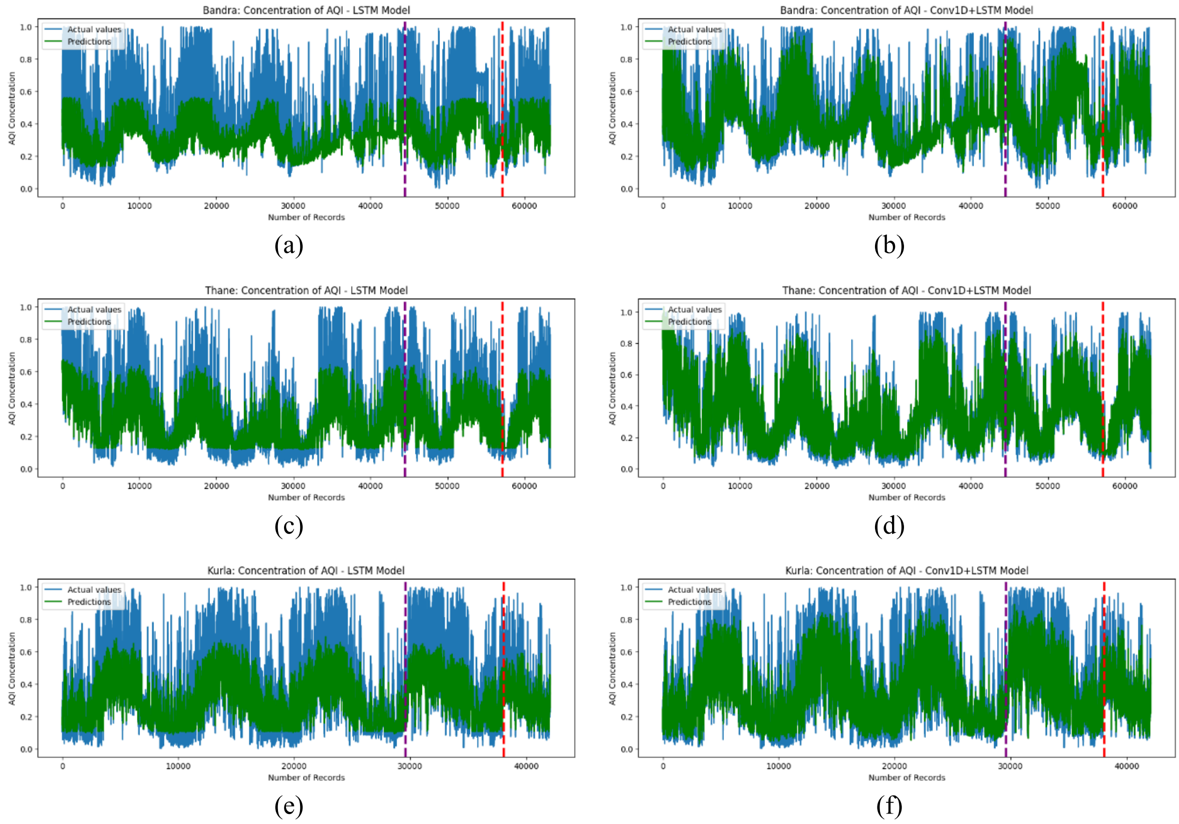


Figure 3. Training and testing forecast of AQI for all six sites: (a) Bandra (LSTM), (b) Bandra (Conv1D+LSTM), (c) Thane (LSTM), (d) Thane (Conv1D+LSTM), (e) Kurla (LSTM), and (f) Kurla (Conv1D+LSTM)

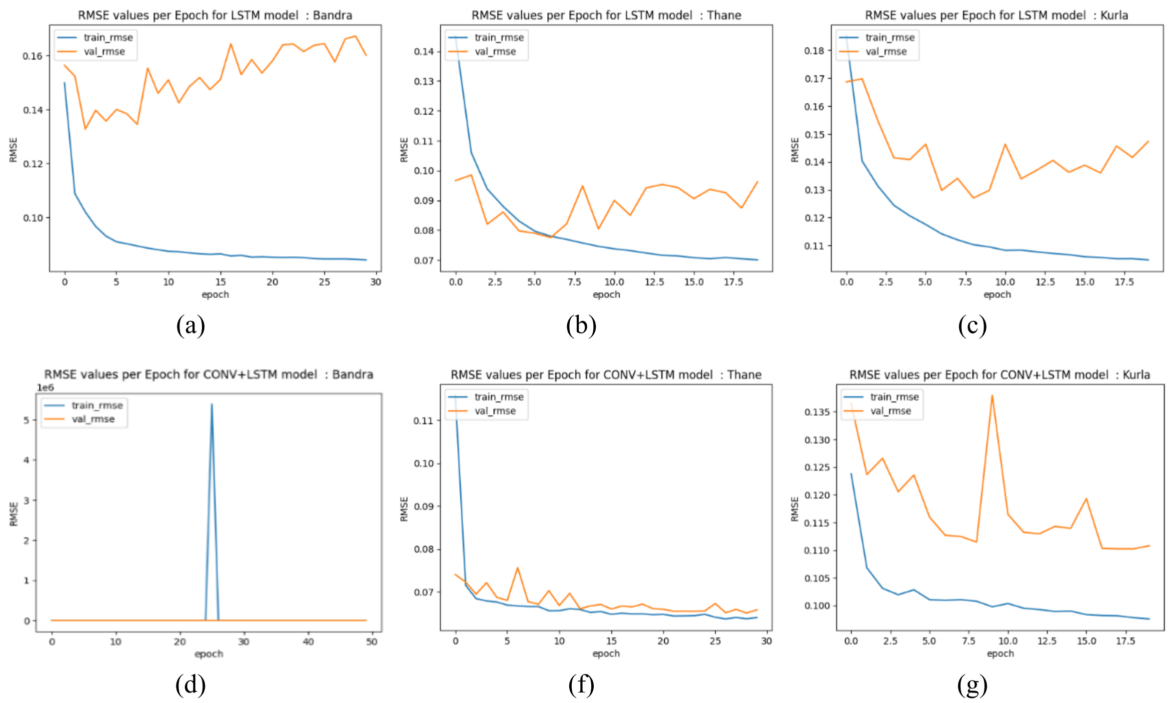


Figure 4. Training history of both LSTM and CPNv1d+LSTM models for all six sites: (a) Bandra (LSTM), (b) Thane (LSTM), (c) Kurla (LSTM), (d) Bandra (Conv1D+LSTM), (e) Thane (Conv1D+LSTM), and (f) Kurla (Conv1D+LSTM)

4. CONCLUSION

Air pollution is a brutal reality, particularly in Asian cities, with India carrying an unfair amount of the burden. A majority of the Indian population is exposed to extremely poor air quality, posing serious health hazards. The goal of this research is to forecast the hourly AQI index for six sites in Mumbai using a comparative analysis of an LSTM model and an integrated proposed hybrid Conv1D+LSTM model. The particle swarm optimizer was used to determine the optimal hyperparameters for all the networks under study.

Based on the evaluation indicators and by performing a comparative analysis across six sites, Bandra, Thane, Mazgaon, Kurla, Nerul and Malad, it was discovered that proposed hybrid Conv1D+LSTM projected AQI values with less deviation from actual AQI values which can be understood from MAPE and RMSE values achieved. The R^2 score also shows that for all sites, the Conv1D+LSTM model could understand the variation in the data more precisely. Specifically, the performance of the hybrid model gave average percentage improvement on RMSE values as 28.72%, average improvement in R^2 Score is up to 293.77%, and average MAPE enhancements was up to 38.24%. Therefore, the use of LSTM units in conjunction with 1D convolutional layers was adequate for capturing both local and global patterns over time. This could be due to the convolution layer smoothing the AQI time series and providing that smoothed data to the LSTM layer to forecast. Additionally, Conv1d layer can act as feature extractor and capture localized pattern which might be overlooked by standalone LSTM layers. It was also noted the training history of LSTM network indicates overfitting whereas Conv1D+LSTM are more robust but still it can be further enhanced by using regularization techniques or data augmentation.

Additional variables, such as population density, traffic density, distance from the coast, landfills can be used to provide even more meaningful information. A substitute for hybrid models could be ensemble techniques, which combine the results of multiple weak models to increase forecast accuracy. For constructing a robust model, high-end deep network models such as autoencoders or transformers can be integrated with statistical time series analysis techniques.

ACKNOWLEDGMENT

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.

REFERENCES

[1] D. Mitra and A. Saha, "IoT-based air pollution detection, monitoring and controlling system," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 25, pp. 2173–2182, 2022, doi: 10.1080/09720529.2022.2133254.
[2] S. L. V and V. M. S., "An intelligent deep learning based AQI prediction model with pooled features," *Journal of Theoretical and Applied Information Technology*, vol. 102, no. 1, 2024.
[3] D. P. Shah and D. P. Patel, "A comparison between national air quality index, India and composite air quality index for Ahmedabad, India," *Environmental Challenges*, vol. 5, p. 100356, Dec. 2021, doi: 10.1016/j.envc.2021.100356.
[4] K. D. Rizwan and A. B. Patankar, "Comparative investigation of machine learning and deep learning methods for univariate AQI forecasting," *International Conference on Intelligent Computing and Networking*, pp. 185–196, 2023.
[5] M. Méndez, M. G. Merayo, and M. Núñez, "Machine learning algorithms to forecast air quality: a survey," *Artificial Intelligence Review*, vol. 56, no. 9, pp. 10031–10066, Sep. 2023, doi: 10.1007/s10462-023-10424-4.
[6] R. Gupta and P. Singla, "Prediction of AQI using hybrid approach in machine learning," *ICTACT Journal on Soft Computing*, vol. 13, no. 3, pp. 2917–2921, Apr. 2023, doi: 10.21917/ijsc.2023.0412.
[7] K. D. Rizwan and A. B. Patankar, "Its meteorology based prediction using LSTM neural network," in *Data Science and Big Data Analytics: Proceedings of IDBA 2023*, 2024, pp. 331–343.
[8] R. Yan, J. Liao, J. Yang, W. Sun, M. Nong, and F. Li, "Multi-hour and multi-site air quality index forecasting in Beijing using CNN, LSTM, CNN-LSTM, and spatiotemporal clustering," *Expert Systems with Applications*, vol. 169, p. 114513, May 2021, doi: 10.1016/j.eswa.2020.114513.
[9] P.-W. Soh, J.-W. Chang, and J.-W. Huang, "Adaptive deep learning-based air quality prediction model using the most relevant spatial-temporal relations," *IEEE Access*, vol. 6, pp. 38186–38199, 2018, doi: 10.1109/ACCESS.2018.2853662.
[10] Z. Wu, W. Zhao, and Y. Lv, "An ensemble LSTM-based AQI forecasting model with decomposition-reconstruction technique via CEEMDAN and fuzzy entropy," *Air Quality, Atmosphere & Health*, vol. 15, no. 12, pp. 2299–2311, Dec. 2022, doi: 10.1007/s11869-022-01252-6.
[11] S. Roy and P. Mukherjee, "Air quality index forecasting using hybrid neural network model with LSTM on AQI," *Proceedings on Engineering Sciences*, vol. 2, no. 4, pp. 431–440, Dec. 2020, doi: 10.24874/PES02.04.010.
[12] H. He and F. Luo, "Study of LSTM air quality index prediction based on forecasting timeliness," *IOP Conference Series: Earth and Environmental Science*, vol. 446, no. 3, p. 032113, Feb. 2020, doi: 10.1088/1755-1315/446/3/032113.
[13] L. Zhoul, M. Chenl, and Q. Ni, "A hybrid prophet-LSTM model for prediction of air quality index," in *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, Dec. 2020, pp. 595–601. doi: 10.1109/SSCI47803.2020.9308543.
[14] R. Xu, D. Wang, J. Li, H. Wan, S. Shen, and X. Guo, "A hybrid deep learning model for air quality prediction based on the time–frequency domain relationship," *Atmosphere*, vol. 14, no. 2, p. 405, Feb. 2023, doi: 10.3390/atmos14020405.
[15] Y. Hua, Z. Zhao, R. Li, X. Chen, Z. Liu, and H. Zhang, "Deep learning with long short-term memory for time series prediction," *IEEE Communications Magazine*, vol. 57, no. 6, pp. 114–119, Jun. 2019, doi: 10.1109/MCOM.2019.1800155.
[16] N. S. Gupta, Y. Mohta, K. Heda, R. Armaan, B. Valarmathi, and G. Arulkumaran, "Prediction of air quality index using machine learning techniques: a comparative analysis," *Journal of Environmental and Public Health*, vol. 2023, pp. 1–26, Jan. 2023,

doi: 10.1155/2023/4916267.
[17] M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer, and P. N. Suganthan, "Ensemble deep learning: a review," *Preprint arXiv:2104.02395*, 2021, doi: 10.48550/arXiv.2104.02395.
[18] M. Z. Alom *et al.*, "A state-of-the-art survey on deep learning theory and architectures," *Electronics*, vol. 8, no. 3, p. 292, Mar. 2019, doi: 10.3390/electronics8030292.
[19] R. T. Wahyuni, D. Hanafi, M. R. Tomari, and D. S. Sihabudin Sahid, "Research trends in spatial modeling of $PM_{2.5}$ concentration using machine learning: a bibliometric review," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 37, no. 2, p. 1317, Feb. 2025, doi: 10.11591/ijeecs.v37.i2.pp1317-1327.
[20] M. Huang, B. Q. Hu, H. Jiang, and B. W. Fang, "A water quality prediction method based on k-nearest-neighbor probability rough sets and PSO-LSTM," *Applied Intelligence*, vol. 53, no. 24, pp. 31106–31128, Dec. 2023, doi: 10.1007/s10489-023-05024-2.
[21] P. Saini and B. Nagpal, "PSO-CNN-Bi-LSTM: a hybrid optimization-enabled deep learning model for smart farming," *Environmental Modeling & Assessment*, vol. 29, no. 3, pp. 517–534, Jun. 2024, doi: 10.1007/s10666-023-09920-2.
[22] CPCBCCR, "India AQI: National air quality index," *CPCBCCR*. https://app.cpcbccr.com/AQI_India (accessed Jun. 24, 2024).
[23] OpenWeather, "OpenWeather Homepage," *OpenWeather*. https://openweathermap.org/ (accessed Apr. 21, 2024).
[24] H. P. Vinutha, B. Poornima, and B. M. Sagar, "Detection of outliers using interquartile range technique from intrusion dataset," in *Information and Decision Sciences*, 2018, pp. 511–518. doi: 10.1007/978-981-10-7563-6_53.
[25] H. S. Salem, M. A. Mead, and G. S. El-Taweel, "Particle swarm optimization-based hyperparameters tuning of machine learning models for big COVID-19 data analysis," *Journal of Computer and Communications*, vol. 12, no. 03, pp. 160–183, 2024, doi: 10.4236/jcc.2024.123010.
[26] D. Schürholz, S. Kubler, and A. Zaslavsky, "Artificial intelligence-enabled context-aware air quality prediction for smart cities," *Journal of Cleaner Production*, vol. 271, p. 121941, Oct. 2020, doi: 10.1016/j.jclepro.2020.121941.
[27] M. Nabanita and S. Tanuja, "Prediction of dew point temperature and relative humidity for Nashik region using LSTM," 2023, pp. 477–489. doi: 10.1007/978-981-99-2710-4_39.
[28] L. Jiao and W. Zheng, "Multi-step traffic flow prediction method based on the Conv1D + LSTM," in *2020 Eighth International Conference on Advanced Cloud and Big Data (CBD)*, Dec. 2020, pp. 113–118. doi: 10.1109/CBD51900.2020.00029.
[29] Q. Wang, K. Kang, Z. Zhang, and D. Cao, "Application of LSTM and CONV1D LSTM network in stock forecasting model," *Artificial Intelligence Advances*, vol. 3, no. 1, pp. 36–43, Apr. 2021, doi: 10.30564/aia.v3i1.2790.
[30] N. Mandal and T. Sarode, "A framework for cloud cover prediction using machine learning with data imputation," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 1, p. 600, Feb. 2024, doi: 10.11591/ijece.v14i1.pp600-607.

BIOGRAPHIES OF AUTHORS

Darakhshan Khan has completed her Bachelor of Engineering (BE) in information technology and Mater of Engineering (ME) in computer engineering from University of Mumbai in 2012 and 2015, respectively. She is currently pursuing her Ph.D. Technology in computer engineering from University of Mumbai. She is currently an UGC approved Assistant Professor at Thadomal Shahani Engineering College, Mumbai. Her research focuses on artificial intelligence, deep learning, algorithms and data structures. She can be contacted at email: darakshan.khan@thadomal.org.

Archana B. Patankar has received her Ph.D. in engineering from Mukesh Patel School of Technology Management and Engineering (MPSTME), NMIMS University in 2011. She is currently an UGC approved Professor at Thadomal Shahani Engineering College, Mumbai. She has more than 110 publications in international journals and conferences. Under her guidance, many students have completed their Ph.D. and many are enrolled. Her research focuses on image and signal processing, security systems, data science, cryptography, and block chain. She can be contacted at email: archana.patankar@thadomal.org.

Jyotika Kakar has completed her BE in computer engineering from Thadomal Shahani Engineering college, University of Mumbai in 2024. She has secured admission for Masters in University of California, San Diego, US. Her research focuses on deep network architectures, satellite image processing, and data science. She can be contacted at email: jyotikakakar@gmail.com.