Assessing Real-Time Health Impacts of outdoor Air Pollution through IoT Integration

Pradeep Mullangi
Department of ECE, Shri Vishnu Engineering College for Women, Bhimavaram, India
pradeepm999@gmail.com

K. M. V. Madan Kumar
Department of CSE, Vignan Institute of Technology and Science, Hyderabad, Telangana State, India
madankukunuri@gmail.com

Gera Vijaya Nirmala
Department of ECE, CVR College of Engineering, Hyderabad, Telangana State, India
vijaya.4b8@gmail.com

Ramesh Chandra Aditya Komperla
Geico, USA
akomperla@gmail.com

Nagalinagam Rajeswaran
Electrical and Electronics Engineering, Malla Reddy College of Engineering, Secunderabad, India
rajeswarann@gmail.com (corresponding author)

Amar Y. Jaffar
Computer and Network Engineering Department, College of Computing, Umm Al-Qura University, Makkah, Saudi Arabia
ayjaafar@uqu.edu.sa

Abdullah Alwabli
Department of Electrical Engineering, College of Engineering and Computing in Alqunfudah, Umm Al-Qura University, Mecca 21955, Saudi Arabia
aswabli@uqu.edu.sa

Saeed Faisal Malky
Department of Electrical and Electronics Engineering, University of Jeddah, Jeddah, Saudi Arabia
sfmalki@uj.edu.sa

Received: 30 January 2024 | Revised: 19 February 2024 and 5 March 2024 | Accepted: 13 March 2024

ABSTRACT

Air pollution constitutes a significant global challenge in both public health and the environment, particularly for countries undergoing industrialization and transitioning from low- to middle-income economies. This study aims to investigate the feasibility and effectiveness of a real-time air quality prediction system based on data collected from Internet of Things (IoT) sensors to help people and public institutions track and manage atmospheric pollution. The primary objective of this study was to investigate whether an IoT-based approach can provide accurate and continuous real-time air quality forecasting. The standard dataset provided by the Indian government was analyzed using regression, traditional Long-Short-Term Memory (LTSM), and bidirectional LSTM (BLSTM) models to evaluate their performance on

multivariate air quality features. The results show that the proposed BLSTM model outperformed the other models in minimizing RMSE errors and avoiding overfitting.

Keywords-Internet of Things; air pollution; LSTM; health controllers

I. INTRODUCTION

As the world faces numerous challenges in the field of public health, the widespread issue of atmospheric pollution remains a central concern due to its pervasive impact and cross-border implications [1-2]. Atmospheric pollution, characterized by the presence of harmful substances, such as particulate matter, nitrogen dioxide, and sulfur dioxide, originates from various sources, including industrial emissions, vehicular exhaust, and biomass burning. These pollutants are linked to a wide range of health problems, from respiratory to cardiovascular diseases and cancers, contributing to an estimated 11.65% of global mortality rates. Despite ongoing efforts to mitigate these effects, traditional monitoring and control mechanisms often lack real-time responsiveness and localized accuracy. In recent decades, the world has experienced significant economic advances and urban growth. However, these developments have also been accompanied by serious environmental challenges, such as air pollution. This escalating environmental crisis has adversely affected public health and exacerbated the severity of climate change [3]. Furthermore, the increasing number of vehicles, industrial operations, and other sources of combustion engines has intensified the environmental health crisis. Government bodies play a key role in mitigating these problems, as they strive to combat air pollution by implementing national and global policies and regulations that promote sustainable growth [4]. However, effective management of these environmental issues requires a unified approach from all participants. This includes government entities, organizations, industrial sectors, corporations, local communities, and individuals. The investigation of air quality has become a focal point in recent years [5]. The World Health Organization (WHO) has identified air pollution as a major factor in global disease proliferation, responsible for more than 7 million premature deaths each year [6-7]. These data underscore the immediate need to reduce exposure to air pollution and protect public health, and, as a consequence, air quality monitoring devices have gained significant attention, as they can measure temperature, humidity, and air pollution levels.

In an era marked by rapid industrialization and urbanization, the rising levels of atmospheric pollution represent an urgent call to action [8]. This study underscores the gravity of the situation, emphasizing the need for comprehensive strategies to mitigate air pollution and its adverse health impacts. This involves stringent regulations on industrial and vehicular emissions, promoting renewable energy sources, and raising public awareness of the health risks associated with poor air quality [9]. The rapid escalation in air pollution, driven by intense industrialization and urban expansion, urgently calls for immediate countermeasures [10]. This study illuminates the critical nature of this issue and advocates for the creation of comprehensive strategies to reduce air pollution and its harmful health consequences. These strategies involve strict control over industrial and vehicular emissions, endorsement of renewable energy sources, and raising public awareness regarding the health risks associated with poor air quality. Indoor air pollution alone accounts for roughly 4.1% of global mortality rates. In 2019, the death toll from indoor air pollution exceeded 600,000. A viable solution to this problem lies in the realm of Internet of Things (IoT) monitoring systems [11]. Equipped with cutting-edge sensor technology, these systems can persistently track and collect data on a variety of indoor or outdoor air quality indicators, such as levels of pollutants like CO_2, CO, particulate matter, and volatile organic compounds. Once these data are collected, they can be transmitted to a cloud-based platform for processing and interpretation. When applying machine learning and predictive analysis algorithms to these data, it becomes feasible to generate instant information on air quality and project future patterns [12].

In recent years, the advent of IoT technologies combined with advanced machine learning models has opened new avenues to address air quality concerns. In [13], sensor data were used to perform time-series prediction on air quality parameters and pollutants by applying linear regression. In [14], a systematic review of deep learning models was conducted for time-series air quality forecasting. In [15], supervised models were used to predict the concentration of air pollutants in multiple locations of a city by using spatial-temporal relationships. In [16], air pollution in India, which is exacerbated by rapid urbanization and transportation development, was investigated. This study noted a discrepancy between current air quality monitoring systems and actual pollutant exposure at ground level, where humans directly inhale vehicle emissions. This study presented a real-time monitoring system using sensors to detect key pollutants from vehicular emissions. Leveraging a deep learning-based LSTM algorithm, the system forecasts pollutant levels, assisting decision-making to improve air quality and allowing citizens to accordingly plan their activities. This study also compared predicted ground-level pollutants with ambient air quality levels, highlighting the system's practical value. In [17], a comparative study of various statistical and deep learning methods was conducted to forecast pollution trends of PM2.5 and PM10. The results showed that auto-regressive and Seasonal Auto-Regressive Integrated Moving Average (SARIMA) outperformed deep-learning methods on a limited dataset in Kolkata, India. In [18], a predictive model for pollutant emissions was presented for an airport, based on the number of takeoff and landing cycles. In [19], various shallow, deep, and hybrid learning models were reviewed to determine their advantages and limitations.

Recent studies have explored various IoT-based models for air quality monitoring, employing technologies ranging from simple sensor networks to complex machine learning algorithms. Although these approaches offer significant information, they often lack real-time prediction capabilities or do not adequately account for localized pollution sources. This research gap underscores the need for a more dynamic and adaptable solution that is capable of providing accurate, real-time air quality predictions. The current study seeks to bridge this gap by introducing a novel IoT framework enhanced with bidirectional LSTM models to offer a promising avenue for comprehensive and timely air quality monitoring. The main objective of this study is to offer a scalable solution that can adapt to varying environmental conditions, thus enhancing public health responses by integrating environmental monitoring with public health initiatives.

II. MATERIALS AND METHODS

A. Datasets

This study used the dataset collected and distributed by the Central Pollution Control Board (CPCB) in conjunction with the Ministry of Environment, Forests, and Climate Change [20]. The specific dataset was created utilizing IoT and cloud technologies. The CPCB maintains an extensive database of pollution levels and closely collaborates with state pollution control boards and other government bodies to enact and monitor environmental legislation and regulations and increase public awareness of environmental issues. To collect real-time air quality data, the CPCB has deployed a network of IoT-enabled sensors at diverse locations. The data collected are stored in a cloud-based time-series database, facilitating access to air quality information across different regions. The publicly accessible real-time Air Quality Index (AQI) data can be used to issue warnings and provide air quality evaluations for various areas. Moreover, these data can be exploited for research purposes and time series analysis. The dataset encapsulates 42,000 entries and 16 variables, which include meteorological attributes, such as Benzene, Toluene, Xylene, PM2.5, PM10, NO_2, NO, NO_X, NH_3, CO, SO_2, O_3, AQI_Bucket, and AQI, along with City and Date. The data, collected on an hourly basis from various stations in different states, span five years, from January 1, 2015, to July 31, 2020. The real-time IoT-enabled dataset, stored in a cloud-based database, offers air quality information for diverse locations and is publicly accessible for research and time series analysis. This study utilized a computing environment powered by an AMD Ryzen 5 4500U@2.38GHz processor with 16GB of RAM. All experiments were run using Jupyter Notebook operating on Windows 11 Pro 64-bit. Python 3.6.5 was employed along with multiple open-source libraries, such as Pandas and NumPy. The setup also incorporated libraries, like Matplotlib, Statsmodel, and Sklearn, to optimize the experimentation process.

B. Methodology

Predicting air quality has become crucial in providing early warnings and managing urban air pollution. The objective is to forecast fluctuations in the PM2.5 air pollution index at specific monitoring points over a given timeframe. The observation period is set to one hour, a standard determined by terrestrial air quality monitoring stations. Figure 2 displays a representative example of air pollution data, such as PM2.5 levels, between 2015 and 2020. The task of predicting PM2.5 concentrations can be defined as follows: Given a certain time T, the objective is to predict the PM2.5 concentration values $P_{i,T+1}$ at time $T+1$ or $P_{i,T+n}$ at time $T+n$, taking into account the historical air-quality time series data $AQD = \{AQD_{i,t} | i \in 0, t = 1,2,3 \dots T\}$ from the past [21]. AQD stands for historical air quality-related data, O represents the total observation points, and AQD encompasses not only PM2.5, but also other air quality-related time series data, such as pressure, temperature, and wind speed.

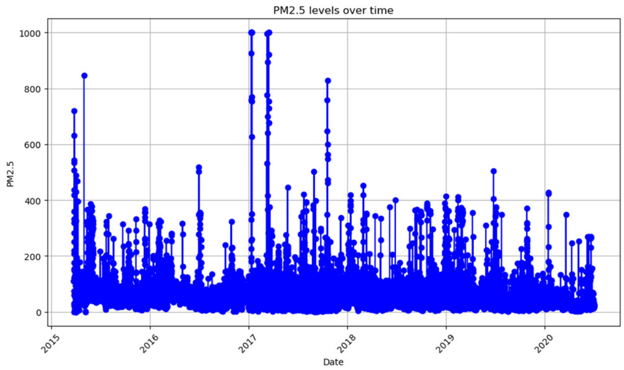

Fig. 1. Air pollution data between 2015 and 2020.

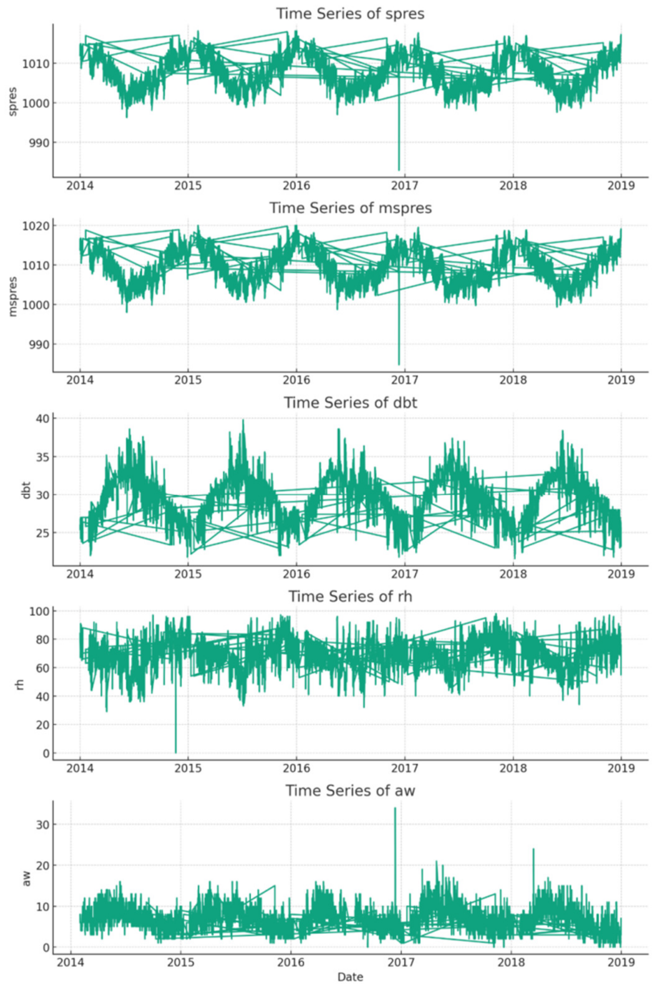

Fig. 2. Average weather pressure, humidity, temperature, and windfall between 2014 to 2019.

As noticed in Figure 2, air quality data typically contain real-valued PM2.5 pollutants, while some datasets may also include CO_2 and PM10. Alongside pollutant data, meteorological observation data play a significant role in determining air quality. For example, high wind speed tends to decrease PM2.5 concentration, elevated humidity often exacerbates air pollution, and high atmospheric pressure generally leads to better air quality. These characteristics are vital for predicting air quality. The crux of air quality prediction lies in how to process and capture the spatial-temporal features of the aforementioned air quality data items. Looking at the PM2.5 data for a month's observation data points, e.g. from 01/01/2010 to 01/31/2010, it is evident that there is contextual information among the observation points in the PM2.5 and wind speed time series. The historical state exerts some influence on the evolution of future trends. In other words, adjacent data points and periodic intervals of air quality time series data typically exhibit a strong correlation with each other.

C. *Overview of the Deep Air Quality Forecasting Framework*

The proposed deep-learning architecture is a modified Bidirectional LSTM (BLSTM) model, designed to capture the spatial-temporal dependencies of air quality-related time series data. Given the correlations among local trend characteristics and the long-term dependencies of multivariate air quality time series data, particularly PM2.5, the time series data are interrelated with other air quality data, and these factors are intrinsically interconnected. Figure 3 offers a visual representation of this deep air quality forecasting framework. This study focuses on the BLSTM for Air Quality Forecasting (BL-AQF) model to process multiple one-dimensional time series data and effectively grasp the spatial-temporal characteristics of various air quality indicators.

The primary phase of this study encompassed sequential training models to extract local trend attributes and potential spatial associations from multistation air quality readings. The model does not analyze the features of each time series in isolation. Instead, it simultaneously processes all the time series data collected from each monitoring point spread across various stations. Subsequently, the extracted features, comprising the local trend features of each station data and the potential spatial correlation features of multistation data, from numerous one-dimensional CNNs are concatenated and fed into a specific BLSTM. This BLSTM learns spatial-temporal dependency features from both past and future contexts, concurrently using time series in both forward and backward directions [22].

The computational components of a typical BLSTM are:

$$i_t = \sigma\left(U^{(i)}x_t + W^{(i)}h_{t-1} + b_i\right) \quad (1)$$

$$f_t = \sigma(U^{(f)}x_t + W^{(f)}h_{t-1} + b_f) \quad (2)$$

$$o_t = \sigma\left(U^{(o)}x_t + W^{(o)}h_{t-1} + b_o\right) \quad (3)$$

$$\tilde{s}_t = tanh\left(U^{(c)}x_t + W^{(c)}h_{t-1} + b_c\right) \quad (4)$$

$$s_t = f_t \cdot s_{t-1} + i_t \cdot \tilde{s}_t \quad (5)$$

$$h_t = o_t \cdot tanh(s_t) \quad (6)$$

In the equations, the input gate determines the influx of new information into the memory cell. The symbol f_t stands for the forget gate, dictating the volume of data to be removed. Meanwhile, o_t serves as the output gate, deciding the measure of data to be relayed either to the subsequent step or directly to the output. $\tilde{s}_t$ functions as a neuron equipped with a self-repeating cell akin to RNNs. The memory cell within the LSTM block denoted s_t, is an aggregation of two components. The initial segment is derived from the prior internal memory state s_{t-1} and the forget gate f_t. The subsequent part is determined by the element-wise product of the self-repeating state $\tilde{s}_t$ and the input gate [23]. A limitation of conventional LSTMs is their ability to harness only the antecedent context of sequential data [24]. In contrast, BLSTM can interpret time series information in a dual-directional manner deploying two distinct hidden layers. Subsequently, the data from both directions are merged and advanced to the output stage. Put differently, BLSTM cyclically assesses time series data in two directions: the forward layer goes from t = 1 to T, while the backward layer operates in reverse form, from $t = T$ to 1.

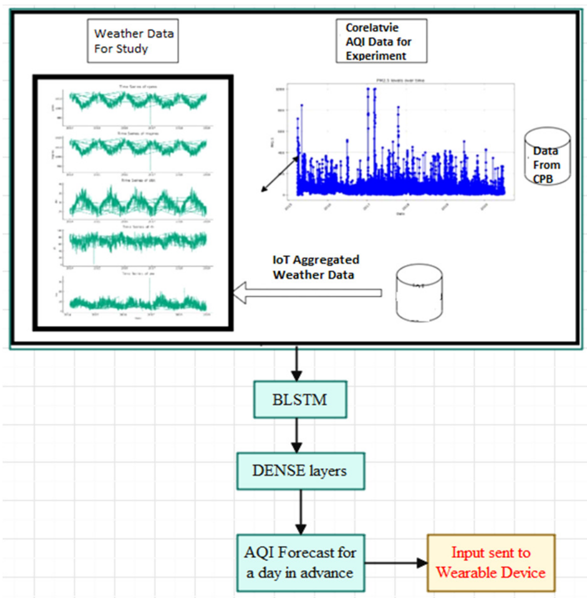

Fig. 3. Architecture of the AQI monitoring system.

III. RESULTS AND DISCUSSION

A. *Understanding the Influence of Traditional Machine Learning in Air Quality Forecasts*

Linear regression, a foundational statistical method, plays a pivotal role in air quality projections by drawing a clear correlation between factors, such as temperature variations, wind dynamics, and moisture levels with air quality index and offering an intuitive way to predict future atmospheric conditions. Its straightforward nature and easy-to-follow logic have cemented its status among environmental analysts. Beyond the linear approach, the Decision Tree (DT) offers a structured method to dissect data based on specific criteria. For example, it can evaluate air quality differences that stem from fluctuating temperatures. This approach excels in highlighting intricate connections and dependencies among various environmental elements. Not only does the tree-like representation bring clarity, but also enhances the comprehensibility of the model logic. Building upon the foundation of DT, Random Forest (RF) introduces an ensemble technique that amalgamates insights from numerous trees. This collective approach mitigates the typical pitfalls of individual trees, such as oversensitivity to data discrepancies. When predicting air quality, it ensures a balanced and comprehensive understanding, harnessing the insights of multiple evaluations for a more refined prediction.

The Mean Squared Error (MSE) of the linear regression was 4.963×10^{-34} while R^2 was 1.0. The MSE for linear regression, being extraordinarily close to zero, and the R^2 value indicate almost perfect predictions that match the actual data. Although this might seem impressive, such perfect scores can be indicative of overfitting, where the model might be too closely tailored to the training data and may not perform as well on new unseen data. For DT, MSE was 2.914×10^{-7} and R^2 was 0.9983. DT had a slightly higher MSE compared to linear regression. However, its MSE is still very low, indicating good prediction accuracy. The R^2 value is slightly less than 1, indicating that the model explains approximately 99.83% of the variance in the dependent variable. This is an excellent score, but the model might be complex and could risk overfitting, given the nature of DT. For RF, MSE was 2.05986×10^{-7} and R^2 was 0.99988. RF, being an ensemble method, combines multiple DTs to produce predictions. Its MSE is lower than that of the DT, exhibiting higher accuracy. The R^2 value is very close to 1, suggesting that the model explains approximately 99.99% of the variance in the dependent variable. This model seems to provide a balanced trade-off between complexity and accuracy.

However deep-learning algorithms are better for AQI prediction due to:

- Complex relationships: Air quality is influenced by a variety of factors, including pollutants, weather patterns, industrial activities, and more. Deep learning can capture intricate, nonlinear relationships among these variables more effectively than shallow algorithms
- Feature extraction: Deep learning models, especially neural networks, can automatically extract and learn important features from raw data, removing the need for manual feature engineering, which can be required for shallow models.
- Handling large datasets: Air quality datasets can be vast, containing data from multiple sensors over long periods. Deep learning models excel when trained on large datasets, allowing them to generalize better.
- Temporal dependencies: RNNs and LSTM networks, subsets of deep learning, are well suited for time-series data, such as air quality indices, as they can remember past information to influence future predictions.
- Regularization techniques: Deep learning models come with a variety of regularization methods, like dropout, which can prevent overfitting, especially when dealing with complex datasets.

B. Results for LSTM-based AQI Prediction

Figure 4 illustrates that RMSE in both the training and validation sets decreases as the number of epochs increases for LSTM. This suggests that the model is learning and refining its predictions with more training. Over the epochs, the RMSE values seem to be stabilizing, indicating that the model is reaching a point of convergence. The gap between the training and validation RMSE appears minimal, suggesting that the model generalizes well and does not overfit the training data. Concerning MSE, both the training and validation MSE values seem to be decreasing as the number of epochs increases. This trend reaffirms that the model is improving its predictions with continued training. Similarly to the RMSE plot, the MSE values appear to stabilize as the epochs progress, suggesting model convergence. Since RMSE is the square root of MSE, the trends observed here are consistent with those seen in the RMSE plot. Considering Mean Absolute Error (MAE), both the training and validation values decrease as the epochs increase, further supporting the assertion that the model is making progress and refining its predictions. As with the previous two metrics, the MAE values seem to level off as the epochs progress, indicating that the model might be reaching an optimal state. The small gap between the training and validation curves suggests good model generalization. All three metrics (RMSE, MSE, MAE) demonstrate a decreasing trend over the epochs, which is a clear sign that the LSTM model is learning and enhancing its predictive capabilities with each training epoch.

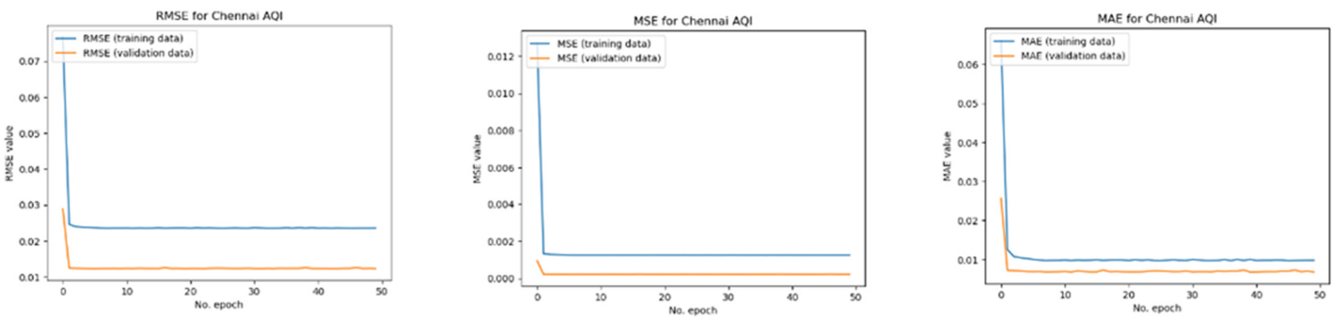

Fig. 4. RMSE, MSE, and MAE results for LSTM-based evaluation.

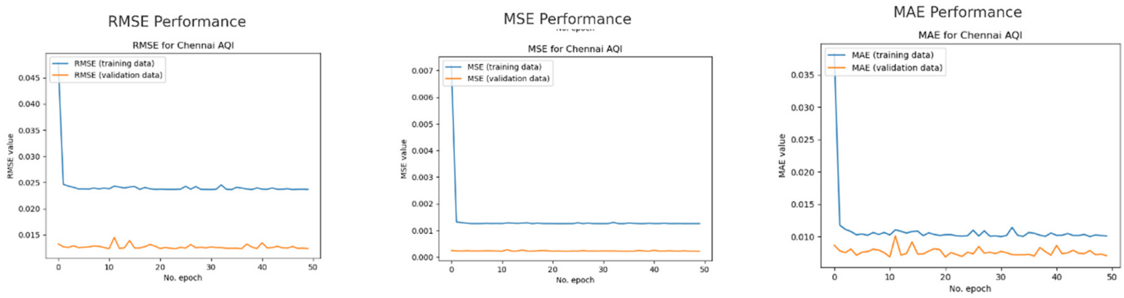

Fig. 5. RMSE, MSE, and MAE of the BLSTM model.

C. Results for BLSTM AQI Prediction

Figure 5 provides a visual representation of the model's performance across epochs in terms of RMSE for both training and validation. Both the training and validation RMSE values rapidly decline during the initial epochs, indicating that the model is learning and improving its prediction accuracy. After the rapid decrease, the RMSE values for both training and validation start to stabilize. This suggests that the model is converging and that further training might not yield significant improvements in prediction accuracy. There is a noticeable gap between the training and validation RMSE values, but it does not appear to significantly widen as training progresses. This is a positive sign, as it suggests that the model is not overfitting the training data. The model shows good generalization, since the validation RMSE is relatively close to the training RMSE, and there is no significant divergence between them over the epochs. Stabilization of the RMSE values indicates that the model may have reached its optimal performance for the given architecture and data. In conclusion, the BLSTM model appears to perform well for multistep AQI forecasting based on the provided RMSE values. Further domain-specific insights and comparisons with actual AQI values would provide a more comprehensive evaluation.

Figure 6 compares the performance of both the BLSTM and the regular LSTM models across epochs. In Epoch 1, BLSTM starts with a training RMSE of 0.1749 and a validation RMSE of 0.0857, while LSTM starts with a training RMSE of 0.2446 and a validation RMSE of 0.1496. BLSTM starts with a better initial performance compared to LSTM. Both models show a rapid decrease in RMSE during the initial epochs, indicating learning and improvements in prediction accuracy. Throughout the epochs, the BLSTM model consistently achieves a lower RMSE than the regular LSTM for both training and validation, showcasing its superior performance. After the initial rapid decrease, the RMSE values for both models begin to stabilize. This indicates that the models reach a point of convergence, where further training might not significantly improve RMSE. BLSTM seems to stabilize at a lower RMSE value compared to LSTM, further emphasizing its efficiency for this task. The gap between the training and validation RMSE for both models is relatively small, suggesting good generalization capabilities on unseen data. However, the BLSTM's validation RMSE remains consistently lower, indicating potentially better generalization for this specific task. BLSTM consistently outperforms regular LSTM in terms of RMSE across all epochs for both training and validation. This can be attributed to its ability to leverage both past and future context in sequence data, making it particularly effective for time series predictions, such as AQI forecasting. Both models demonstrate a converging trend as training progresses, with BLSTM demonstrating a more favorable trajectory. Both models exhibit good generalization, but the BLSTM shows a slight edge in this dataset, given its consistently lower validation RMSE. Overall, for air quality prediction, BLSTM appears to be a more suitable choice than regular LSTM, on the basis of the RMSE values provided.

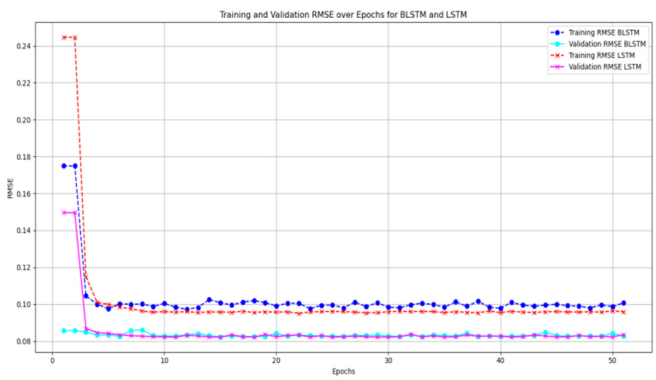

Fig. 6. Comparative performance of LSTM and BLSTM for AQI prediction.

Figure 7 reveals that the actual and predicted values appear to be in close alignment for most of the data points. This suggests that the BLSTM model has captured the underlying patterns and trends in the data quite well. There are certain points where the predicted values deviate from the actual values. These deviations, however, appear to be minimal and not too frequent, indicating that the model performance is generally consistent. The model seems to capture both the peaks and troughs of the actual data, suggesting that it has learned the seasonality or cyclic behavior, if any, present in the dataset. In many sections of the chart, the actual and predicted curves overlap, which is a strong indicator of the model's accuracy in those intervals. BLSTM appears to perform quite effectively for the AQI prediction task, as the close alignment between the actual and predicted values signifies a high degree of accuracy. The ability of the model to capture major trends, peaks, and troughs in the data emphasizes its robustness.

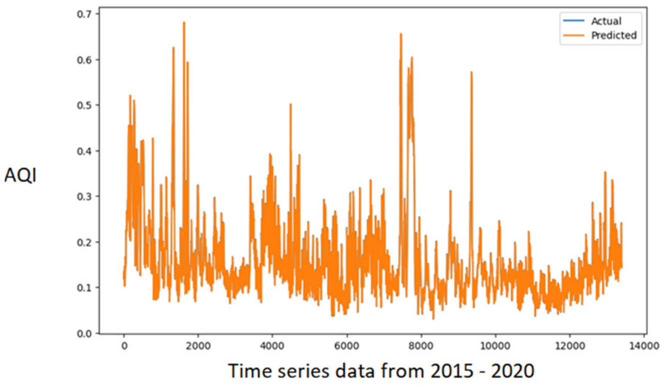

Fig. 7. Actual vs predicted time series data results by BLSTM.

IV. CONCLUSIONS

This study introduces a groundbreaking IoT-based framework that uses BLSTM for precise prediction of air quality. The findings demonstrate a significant improvement in prediction accuracy over traditional methods, avoiding overfitting and underscoring the potential of integrating advanced machine learning techniques with IoT technologies for air pollution monitoring. The application of this study extends beyond academic interest, offering tangible benefits for urban planning, public health strategies, and environmental protection initiatives. By providing accurate real-time air quality data, the proposed BLTSM model supports informed decision-making, allowing timely interventions that can mitigate the adverse health impacts of air pollution and improve the quality of life in urban environments. Looking ahead, the integration of the proposed IoT framework with smart city infrastructure represents an exciting avenue for research and development. Such advances could lead to the creation of highly adaptive and responsive environmental monitoring systems, capable of predicting and managing air quality issues more effectively. Furthermore, exploring the model's performance across a broader range of environmental conditions will be critical to ensure its applicability and reliability on a global scale. The limitations of this study are particularly based on the need for extensive data collection and validation in diverse environmental settings. Future research should aim to address these challenges, seeking to refine and expand the model's capabilities to ensure its effectiveness and scalability. In conclusion, this research contributes a novel approach to the challenge of air quality prediction, highlighting the synergy between IoT technologies and machine learning. Continuous advances in this field grow the potential for impactful real-world applications, offering a beacon of hope for addressing the global issue of air pollution.

REFERENCES

[1] N. Harkat, A. Rahmane, and I. Bendjemila, "The Impact of Industrial Air Pollution on the Urban Environment of Setif: Modeling and Mapping of Total Suspended Particles," *Engineering, Technology & Applied Science Research*, vol. 12, no. 6, pp. 9431–9439, Dec. 2022, https://doi.org/10.48084/etasr.5256.

[2] H. Ritchie and M. Roser, "Indoor Air Pollution," *Our World in Data*, Mar. 2024, [Online]. Available: https://ourworldindata.org/indoor-air-pollution.

[3] E. Fazakas, I. A. Neamtiu, and E. S. Gurzau, "Health effects of air pollutant mixtures (volatile organic compounds, particulate matter, sulfur and nitrogen oxides) – a review of the literature," *Reviews on Environmental Health*, Mar. 2023, https://doi.org/10.1515/reveh-2022-0252.

[4] C. Li *et al.*, "The role of mental health professionals in the climate crisis: an urgent call to action," *International Review of Psychiatry*, vol. 34, no. 5, pp. 563–570, Jul. 2022, https://doi.org/10.1080/09540261.2022.2097005.

[5] H. Wu, Y. Li, Y. Hao, S. Ren, and P. Zhang, "Environmental decentralization, local government competition, and regional green development: Evidence from China," *Science of The Total Environment*, vol. 708, Mar. 2020, Art. no. 135085, https://doi.org/10.1016/j.scitotenv.2019.135085.

[6] F. Ekici, G. Orhan, Ö. Gümüş, and A. B. Bahce, "A policy on the externality problem and solution suggestions in air transportation: The environment and sustainability," *Energy*, vol. 258, Nov. 2022, Art. no. 124827, https://doi.org/10.1016/j.energy.2022.124827.

[7] M. O. Raimi, Z. Adio, O. O. Emmanuel, Timothy Kayode Samson, B. S. Ajayi, and T. J. Ogunleye, "Impact of Sawmill Industry on Ambient Air Quality: A Case Study of Ilorin Metropolis, Kwara State, Nigeria," *Energy and Earth Science*, vol. 3, no. 1, 2020, https://doi.org/10.2139/ssrn.3586971.

[8] D. L. Mendoza *et al.*, "Air Quality and Behavioral Impacts of Anti-Idling Campaigns in School Drop-Off Zones," *Atmosphere*, vol. 13, no. 5, May 2022, Art. no. 706, https://doi.org/10.3390/atmos13050706.

[9] X. Zhang *et al.*, "Linking urbanization and air quality together: A review and a perspective on the future sustainable urban development," *Journal of Cleaner Production*, vol. 346, Apr. 2022, Art. no. 130988, https://doi.org/10.1016/j.jclepro.2022.130988.

[10] P. Kokate, S. Sadistap, and A. Middey, "Atmospheric CO2 Level Measurement and Discomfort Index Calculation with the use of Low-Cost Drones," *Engineering, Technology & Applied Science Research*, vol. 13, no. 5, pp. 11728–11734, Oct. 2023, https://doi.org/10.48084/etasr.6230.

[11] G. Marques and R. Pitarma, "mHealth: Indoor Environmental Quality Measuring System for Enhanced Health and Well-Being Based on Internet of Things," *Journal of Sensor and Actuator Networks*, vol. 8, no. 3, Sep. 2019, Art. no. 43, https://doi.org/10.3390/jsan8030043.

[12] Q. Guo *et al.*, "Applications of artificial intelligence in the field of air pollution: A bibliometric analysis," *Frontiers in Public Health*, vol. 10, Sep. 2022, https://doi.org/10.3389/fpubh.2022.933665.

[13] R. Kumar, P. Kumar, and Y. Kumar, "Time Series Data Prediction using IoT and Machine Learning Technique," *Procedia Computer Science*, vol. 167, pp. 373–381, Jan. 2020, https://doi.org/10.1016/j.procs.2020.03.240.

[14] N. Zaini, L. W. Ean, A. N. Ahmed, and M. A. Malek, "A systematic literature review of deep learning neural network for time series air quality forecasting," *Environmental Science and Pollution Research*, vol. 29, no. 4, pp. 4958–4990, Jan. 2022, https://doi.org/10.1007/s11356-021-17442-1.

[15] A. Gilik, A. S. Ogrenci, and A. Ozmen, "Air quality prediction using CNN+LSTM-based hybrid deep learning architecture," *Environmental Science and Pollution Research*, vol. 29, no. 8, pp. 11920–11938, Feb. 2022, https://doi.org/10.1007/s11356-021-16227-w.

[16] J. Srishtishree, S. Mohana Kumar, and C. Shetty, "Air Quality Monitoring with IoT and Prediction Model using Data Analytics," in *Innovations in Computer Science and Engineering: Proceedings of 7th ICICSE*, H. S. Saini, R. Sayal, R. Buyya, and G. Aliseri, Eds. Singapore: Springer, 2020, pp. 535–544.

[17] P. Nath, P. Saha, A. I. Middya, and S. Roy, "Long-term time-series pollution forecast using statistical and deep learning methods," *Neural Computing and Applications*, vol. 33, no. 19, pp. 12551–12570, Oct. 2021, https://doi.org/10.1007/s00521-021-05901-2.

[18] D. A. Pamplona and C. J. P. Alves, "Civil Aircraft Emissions Study and Pollutant Forecasting at a Brazilian Airport," *Engineering, Technology & Applied Science Research*, vol. 10, no. 1, pp. 5217–5220, Feb. 2020, https://doi.org/10.48084/etasr.3227.

[19] H. Liu, G. Yan, Z. Duan, and C. Chen, "Intelligent modeling strategies for forecasting air quality time series: A review," *Applied Soft Computing*, vol. 102, Apr. 2021, Art. no. 106957, https://doi.org/10.1016/j.asoc.2020.106957.

[20] M. Mondal, "Air Quality Index." [Online]. Available: https://kaggle.com/code/mahadevmondal/air-quality-index.

[21] Q. B. Jamali *et al.*, "Analysis of CO2, CO, NO, NO2, and PM Particulates of a Diesel Engine Exhaust," *Engineering, Technology & Applied Science Research*, vol. 9, no. 6, pp. 4912–4916, Dec. 2019, https://doi.org/10.48084/etasr.3093.

[22] Z. Zhang and M. Tao, "Deep Learning for Wireless Coded Caching With Unknown and Time-Variant Content Popularity," *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 1152–1163, Oct. 2021, https://doi.org/10.1109/TWC.2020.3030973.

[23] M. S. Balamurugan and R. Manojkumar, "Study of short term rain forecasting using machine learning based approach," *Wireless Networks*, vol. 27, no. 8, pp. 5429–5434, Nov. 2021, https://doi.org/10.1007/s11276-019-02168-3.

[24] C. Zhang *et al.*, "Satellite UV-Vis spectroscopy: implications for air quality trends and their driving forces in China during 2005–2017," *Light: Science & Applications*, vol. 8, no. 1, Nov. 2019, Art. no. 100, https://doi.org/10.1038/s41377-019-0210-6.