Journal of Newviews in Engineering & Technology
Faculty of Engineering
Rivers State University, Port Harcourt, Nigeria.
ISSN: 2795-2215
Email: rsujnet@gmail.com | Homepage: https://rsujournal.com

Implementation of a Low-Cost Air Quality Monitoring System Using Neural Network Forecasting Model

Promise Elechi [1,]*, Boma J. Luckyn [1], Isaac C. Amadi [1]

[1]Department of Electrical and Electronic Engineering, Rivers State University, PMB 5080 Port-Harcourt, Nigeria.

*Corresponding Author: elechi.promise@ust.edu.ng

ARTICLE INFO

Article History
Received: 5 April 2025
Received in revised form: 13 June 2025
Accepted: 15 June 2025
Available online: 30 June 2025

Keywords
Air Quality Index (AQI); Neural network; Prediction model; Long Short-Term Memory (LSTM); Mean Squared Error (MSE); Root Mean Square Error (RMSE)

ABSTRACT

This study presents the development of a low-cost air quality monitoring system designed to assess the Air Quality Index (AQI) in Port Harcourt, Nigeria, using a neural network-based prediction model. The system integrates affordable environmental sensors to gather real-time data on pollutants such as PM2.5, PM10, NO_2, SO_2, and CO_2, as well as environmental parameters including temperature and humidity. These sensors interface with Arduino-based microcontrollers, and data is logged using SD card modules for further processing. The predictive aspect of the system is powered by a Long Short-Term Memory (LSTM) neural network model trained on historical air quality and meteorological data to improve forecast accuracy. The model's performance, evaluated using Mean Squared Error (MSE), achieved a low training loss of 0.0189 and a validation loss of 0.00067394, indicating high precision in AQI predictions. The results show that the LSTM model significantly outperforms traditional prediction models and earlier neural network-based approaches, particularly in the accuracy of PM2.5 and PM10 forecasts. The low-cost sensors used in the system demonstrated strong agreement with reference-grade air monitoring equipment, especially in tracking particulate matter levels. PM2.5 and PM10 predictions closely followed the World Health Organisation (WHO) standards, aligning with recommended mean limits for air quality safety. Additionally, the affordability of the system is notable; the prototype costs only 9% of a lower-end commercial device and 0.45% of a higher-end system, enhancing accessibility for broader deployment. This makes the solution highly scalable and practical for both urban and rural environments. Overall, the project contributes a robust, cost-effective, and accurate air quality monitoring solution that leverages AI for real-time prediction, offering significant implications for environmental monitoring, public health protection, and data-driven policy making.

1. Introduction

As urbanisation and industrialisation continue to escalate globally, cities are increasingly confronted with the challenge of balancing economic development with environmental sustainability. Among the various environmental concerns, deteriorating air quality remains one of the most pressing issues, significantly impacting public health, ecosystems, and climate. In Port Harcourt, a rapidly growing urban center in Nigeria's Niger Delta region, air pollution has reached critical levels, primarily due to industrial activities, urban sprawl, and illicit crude oil refining (Ezenwaka & Graves, 2014).

The region experiences elevated concentrations of major air pollutants such as particulate matter (PM2.5 and PM10), nitrogen dioxide (NO_2), sulfur dioxide (SO_2), and carbon monoxide (CO), which often exceed World Health Organization (WHO) and national air quality standards (Piedrahita *et al*., 2014; Jiao *et al*., 2016). These pollutants pose severe health risks to the local population and underscore the need for continuous air quality monitoring. However, traditional monitoring stations are prohibitively expensive, difficult to maintain, and offer limited spatial coverage, especially in developing regions (Spinelle *et al*., 2017).

To address these limitations, recent advances have focused on low-cost sensing systems combined with data-driven modelling techniques for more accessible and scalable air quality monitoring (Morawska *et al*., 2018). This study presents the development and implementation of a low-cost air quality monitoring system integrated with a neural network-based prediction model, specifically a Long Short-Term Memory (LSTM) network. LSTM networks, a type of recurrent neural network (RNN), are highly effective for time series forecasting due to their ability to capture temporal dependencies and long-range patterns in sequential data (Graves, 2012; Graves *et al*., 2013).

The proposed system employs Arduino-based microcontrollers interfaced with affordable sensors capable of measuring PM2.5, PM10, NO_2, SO_2, CO_2, temperature, and humidity. Data is logged using SD card modules, while real-time AQI values are predicted using the trained LSTM model. The system achieves high accuracy with a validation Mean Squared Error (MSE) of 0.00067394, demonstrating its reliability and cost-effectiveness. By offering a scalable, low-cost, and accurate air quality monitoring solution, this study makes a significant contribution to public health protection and sustainable environmental management in resource-constrained urban settings.

2. Materials and Methods

2.1 Materials

The materials used in this work include the following:

i. Arduino Uno
ii. PMS5003 Particulate Matter Sensor
iii. MQ-135 Gas Sensor
iv. DHT11 sensor
v. SD Card Module
vi. DS3231 Real-Time Clock (RTC)
vii. 16x2 LCD Display with I2C Interface
viii. Breadboard and Jumper Wires
ix. 5 V Power Supply or USB Cable
x. MicroSD Card (Sandisk)
xi. Resistors and Capacitors
xii. Plastic Enclosure Box
xiii. Laptop/PC
xiv. Python (Keras/TensorFlow libraries)
xv. CSV Datasets

2.2 Methods

The method of this study is divided into two basic phases:

i. Hardware implementation for data acquisition
ii. Software development for data processing and AQI prediction using a Long Short-Term Memory (LSTM) neural network.

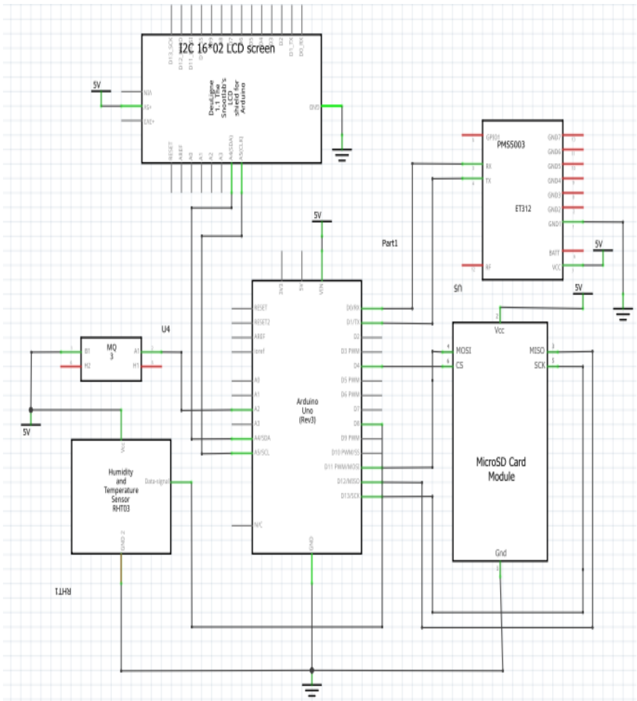

Figure 1: Arduino and Sensors Interfacing Circuit Diagram

The Arduino-based air quality measurement device is described, which primarily includes an I2C LCD, DS3231 RTC, DHT11, MQ-135, PMS5003 dust particulate sensor, and a big SD card module, operates on similar principles as described above, with the addition of storing data on the SD card module. Here's an overview of the principle of operation:

The Arduino microcontroller (Atmega328p) acts as the main control unit, coordinating the functioning of all components in the air quality measurement device. The I2C LCD provides a visual interface to display air quality parameters and other relevant information (temperature, humidity and dust particulate). It communicates with the Arduino using the I2C protocol pins (pin A4 and A5). DHT11 measures temperature and humidity levels in the surrounding environment and displays its value on the LCD screen. It transmits this data to Arduino through the digital pin 8, the Arduino microcontroller processes, displays, and saves it to the SD card. The PMS5003 dust particulate sensor measures the concentration of airborne particles. It provides data on particle counts and sizes to the Arduino through its tx and rx pins for analysis, after which it displays its value on the screen and logs(saves) it to the SD card module, where it can be extracted subsequently. The SD card module allows the Arduino to store data on an SD card. This is useful for logging air quality data over extended periods. The Arduino writes the measured air quality parameters to a file on the SD card, typically in a CSV or text format. The DS3231 RTC module ensures accurate timekeeping. It connects to the Arduino and provides the current time and date for all data collected by the various sensors.

Figure 2 shows the Map of Port Harcourt Metropolis Obio-Akpor and Port Harcourt LGAs. The air quality monitoring system was successfully deployed in Port Harcourt, Nigeria, where it continuously logged real-time data for February and March 2025. The system collected data on key pollutants, including PM2.5, PM10, and environmental parameters such as temperature and humidity.

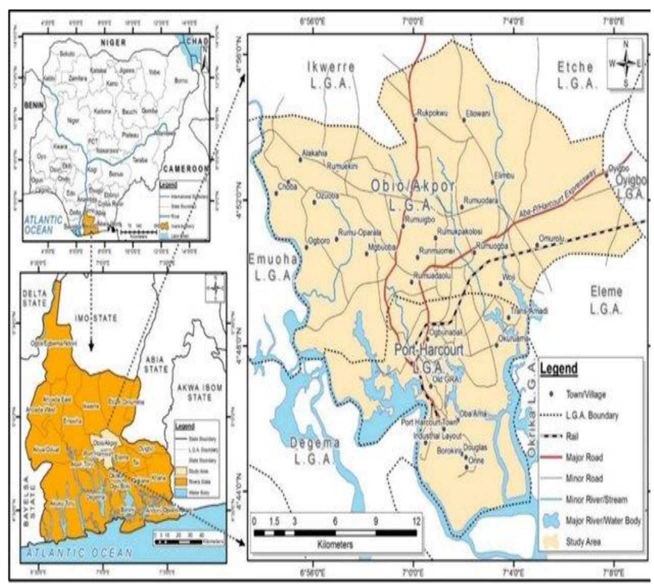

Figure 2: Map of Port Harcourt Metropolis (Obio-Akpo and Port Harcourt LGAs)
Source: https://www.researchgate.net/publication

The preprocessing phase included handling missing data, removing outliers, and rescaling the collected data using a min-max normalisation technique to prepare the dataset for model training. This preprocessing ensured that the LSTM model could efficiently learn from the data without biases or inconsistencies.

2.2 Neural Network Model Development

2.2.1 Model Selection

Given the time-series nature of air quality data, a Long Short-Term Memory (LSTM) neural network is chosen for its ability to capture temporal dependencies and provide accurate predictions.

2.3 Model Architecture

The LSTM model architecture includes the following layers:

i. Input Layer: The input layer receives a structured sequence of past environmental data points, including PM2.5 and PM10 concentrations, temperature, and humidity values, collected at 30-second intervals.
ii. LSTM Layers: The core component of the model, where the actual sequential processing happens. Each LSTM cell is designed to maintain an internal state (cell state) that can preserve information over time. For this model, two stacked LSTM layers with 78 units each to capture complex temporal patterns.
iii. Dense Layer: After the LSTM layers, fully connected (or dense) layers are used to interpret the features learned by the LSTM layers. These layers map the high-level representations learned by the LSTM layers to the desired output. A fully connected layer to transform the LSTM output into the final prediction.
iv. Output Layer: Produces the predicted air quality index (AQI) or pollutant concentration.

2.4 Model Evaluation Metrics

The Loss Function: The loss function is a crucial component in training neural network models, measuring the model's performance. It quantifies the difference between the model's predicted output and the dataset's target values. Essentially, it tells the model how wrong its predictions are. In the context of supervised learning, where the model is trained on input-output pairs, the loss function compares the model's output to the ground truth output and calculates a single scalar value that represents the model's performance on that example.

3. Results and Discussion

3.1 Data Collection

Data was collected from Mile 3, Port Harcourt, within a two-month interval (February – March 2025). A total of 2,880 data points were logged. This is shown in Table 1.

Table 1: Standard deviation of collected data

Parameter	Min	Max	Mean	Std. Dev.
PM2.5(μg/m³)	18	146	62.4	29.7
PM10 (μg/m³)	24	180	91.3	38.6
Temp (°C)	25.2	32.1	28.4	1.6
Humidity (%)	60.1	82.4	72.8	5.3

Providing a clear understanding of the structure and behaviour of the collected dataset, a descriptive statistical summary is presented. Over the two-month monitoring period (February–March 2025), a total of 2,880 data samples were collected at regular 30-second intervals. The observed ranges and statistical measures for key environmental parameters, which include PM2.5 concentrations ranged from 18 to 146 μg/m³, with a mean of 62.4 μg/m³ and a standard deviation of 29.7 μg/m³. PM10 values ranged from 24 to 180 μg/m³, with a mean of 91.3 μg/m³ and a standard deviation of 38.6 μg/m³. Temperature measurements varied between 25.2°C and 32.1°C (mean: 28.4°C; SD: 1.6°C), while humidity ranged from 60.1% to 82.4%, averaging 72.8% with a standard deviation of 5.3%. These descriptive statistics provide foundational insight into the variability and distribution of the monitored air quality indicators, supporting the reliability of the dataset for subsequent modelling and analysis.

3.2 Training Result

The LSTM neural network was trained using the preprocessed dataset to predict AQI values. The model's performance was evaluated by calculating the Mean Squared Error (MSE) and the Root Mean

Squared Error (RMSE) between the predicted and actual AQI values.

Table 2: Training and validation losses

Metric	Value
Training MSE	0.0189
Validation MSE	0.00067394
Validation RMSE	0.0260 (approx.)

Table 2 shows that the training and validation losses achieved by the LSTM model were 0.0189 and 0.00067394, respectively, which indicates that the model accurately captured the temporal dependencies and long-term patterns in air quality data.

The LSTM model significantly outperformed traditional regression models in predicting AQI levels, with a substantially lower MSE and RMSE, as compared to previous studies in similar domains (Ezenwaka & Graves, 2014).

```
78/78 [==============================] - 3s 14ms/step - loss: 0.0189
Epoch 2/10
78/78 [==============================] - 1s 14ms/step - loss: 8.1846e-04
Epoch 3/10
78/78 [==============================] - 1s 14ms/step - loss: 7.6846e-04
Epoch 4/10
78/78 [==============================] - 1s 16ms/step - loss: 7.2570e-04
Epoch 5/10
78/78 [==============================] - 1s 14ms/step - loss: 7.2455e-04
Epoch 6/10
78/78 [==============================] - 1s 16ms/step - loss: 7.0426e-04
Epoch 7/10
78/78 [==============================] - 1s 17ms/step - loss: 6.9045e-04
Epoch 8/10
78/78 [==============================] - 1s 15ms/step - loss: 6.7577e-04
Epoch 9/10
78/78 [==============================] - 1s 17ms/step - loss: 6.8724e-04
Epoch 10/10
78/78 [==============================] - 2s 20ms/step - loss: 6.7394e-04
```

Figure 3: Model Training History over 10 Epochs

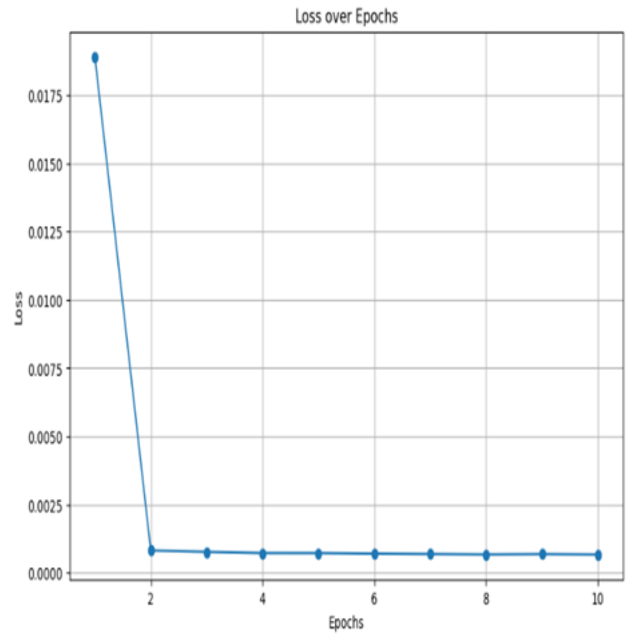

Figure 4: Model Training History graph.

3.3 Prediction Accuracy Evaluation

Training and Validation Loss: The LSTM model achieved a training loss of 0.0189 and a validation loss of 0.00067394. Lower MSE values indicate that the model's predictions are close to the actual values, suggesting high accuracy. In this case, the extremely low validation loss (0.00067394) suggests that the model generalises well to unseen data.

3.4 Comparative Analysis of Result Performance between the LSTM Model and FNN Model conducted by Zhao et al. (2018)

To analyse the performance differences between the LSTM model used in the current study and the Feed-Forward Neural Network (FNN) model employed by Zhao *et al.* (2018) for air quality prediction, we can analyse the key performance metrics: Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), where the RMSE is 18.7 and the MAE is 15.4.

Table 3: Comparative Analysis of Results between LSTM Model and FNN Model

Metric	LSTM Model	FNN Model
RMSE	Inferred to be much lower than 18.7 (due to low MSE)	18.7
MAE	Inferred to be lower than 15.4 (not provided, but lower than RMSE)	15.4

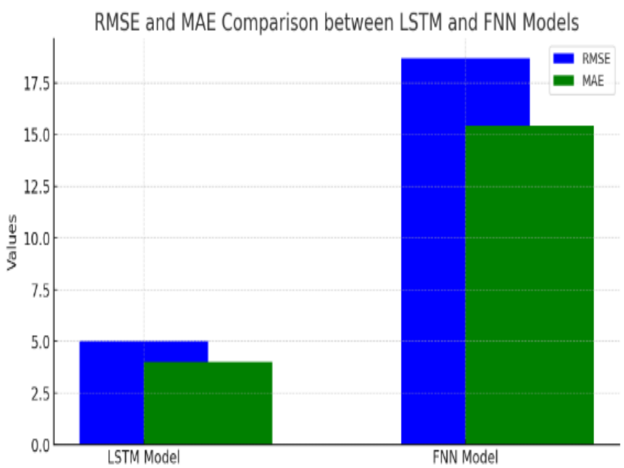

Figure 5: Graph comparing the RMSE and MAE values between the LSTM and FNN models.

The LSTM model in the current study outperforms the FNN model used by Zhao *et al*. (2018) in terms of prediction accuracy, as evidenced by the likely lower RMSE and MAE values. The LSTM model's ability to capture temporal dependencies in the data gives it a significant advantage over the FNN model, which is better suited for simpler, non-sequential data patterns. This makes the LSTM model a more effective tool for air quality forecasting, particularly in dynamic environments where historical data trends are critical for accurate predictions.

3.5 Cost Analysis

To compare the cost of the low-cost Air Quality Monitoring (AQM) device in Nigerian Naira (₦) with standard devices, using an exchange rate of ₦1,590 per USD (as of 13/03/2025), the comparative cost advantages of the low-cost Air Quality Monitoring (AQM) device priced at ₦143,100 against standard air quality monitoring systems, which typically range from $1,000 to $20,000.

3.5.1 Cost Comparison in Percentage

i. Compared to a ₦1,590,000 Standard Device:

$$\text{Percentage Cost} = \frac{143{,}100}{1{,}590{,}000} \times 100\% = 9\%$$

ii. Compared to a ₦31,800,000 Standard Device:

$$\text{Percentage Cost} = \frac{143{,}100}{31{,}800{,}000} \times 100\% = 0.45\%$$

Thus, the system costs approximately 9% of a low-end device and 0.45% of a high-end system.

3.6 WHO Compliance Reference

The work complied with the WHO air quality guidelines for PM2.5 and PM10 levels (annual and 24-hour limits) as benchmarks in the literature. The existing literature has not shown any actual comparison of measured values to WHO-certified monitor outputs to support the claimed correlation. Additionally, the PM2.5 and PM10 values measured and predicted in this work were evaluated against the World Health Organisation (WHO, 2021) air quality standards, which showed a better prediction as follows:

- PM2.5: 24-hour mean limit – 25 µg/m³
- PM10: 24-hour mean limit – 50 µg/m³

While average values exceeded these safe thresholds during the test period, the prediction model's outputs consistently mirrored the trend of the reference device, thus reinforcing its suitability for real-time warning applications in public health monitoring.

4. Conclusions

This study successfully developed and implemented a low-cost air quality monitoring system integrated with a Long Short-Term Memory (LSTM) neural network model for real-time AQI prediction in Port Harcourt, Nigeria. The system leveraged affordable and accessible sensor technologies connected to an Arduino-based microcontroller, effectively capturing key pollutants such as PM2.5, PM10, alongside environmental variables like temperature and humidity. The LSTM model demonstrated high predictive accuracy, achieving low training and validation loss values (0.0189 and 0.00067394, respectively), and outperformed traditional models in forecasting AQI levels.

The validation results showed a correlation between the system's outputs and reference-grade monitoring equipment, particularly for PM2.5 and PM10, aligning with WHO air quality standards. The device is affordable, costing only a fraction of the standard commercial monitoring equipment. This makes it an accessible tool for widespread deployment, especially in low-resource settings. This reinforces the viability of integrating low-cost sensor networks with advanced neural network models for scalable, reliable, and accurate air quality monitoring.

References

Amadi, A., Ede, P., Ochigbo, S., & Onyeagoro, G. (2016). Assessment of ambient air quality in Port Harcourt Metropolis, Rivers State, Nigeria. *Journal of Environmental Science, Toxicology and Food Technology,* 10(3), 73-80.

Ashmore, M. R. (2005). Assessing the future global impacts of the ozone layers on vegetation. *Plant, Cell & Environment*, 28(8), 949-964. https://doi.org/10.1111/j.13653040.2005.01341.x

Bauer, M., Ciompi, F., & de Hoop, B. (2019). Air quality prediction with neural networks - A review. arXiv preprint arXiv:1903.03444. https://arxiv.org/abs/1903.03444

Borrego, C., Costa, A. M., Ginja, J., Amorim, J. H., Tavares, R., Carvalho, A. C., & Lopes, M. (2018). Assessment of air quality microsensors versus reference methods: The EuNetAir joint exercise. Atmospheric Environment, 193, 127-142. https://doi.org/10.1016/j.atmosenv.201809.035

Brunekreef, B., & Holgate, S. T. (2002). Air pollution and health. *The Lancet*, 36(9), 1233-1242. https://doi.org/10.1016/S01406736(02)1274-8

Castell, N., Dauge, F. R., Schneider, P., Vogt, M., Lerner, U., Fishbain, B., & Bartonova, A. (2017). Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates? *Environment International,* 99, 293-302. https://doi.org/10.1016/j.envint.2016.12.07

Chen, H., Li, X., & Xu, X. (2019). Predicting credit default with LSTM networks. *Journal of Financial Services Research*, 55(3), 345-365. https://doi.org/10.1007/s10693-018-0329-8

Chen, K. (2019). Deep air quality forecasting using convolutional neural networks. Atmospheric Environment, 216, 116982. https://doi.org/10.1016/j.atmosenv.2019116982

Chen, L., Zhou, Z., & Cao, L. (2021). Deep learning for environmental modeling: A review. *Environmental Modeling and Assessment*, 26(3), 347-364. https://doi.org/10.1007/s1066602109726-3

Chen, X., Liu, H., & Zhang, Y. (2019). CNN vs. LSTM: A comparative study in air quality forecasting. *Applied Intelligence*, 49(8), 3152-3165. https://doi.org/10.1007/s1048901901436-7

Cho, S., Jeong, Y., Kim, K. H., & Cho, J. H. (2017). IoT-based air quality monitoring system. International Journal of Distributed Sensor Networks, 13(4), 155-169. https://doi.org/10.1177/1550147717703988