Engineering and Technology Journal e-ISSN: 2456-3358
Volume 10 Issue 10 October-2025, Page No.-7665-7673
DOI: 10.47191/etj/v10i10.41, I.F. – 8.482

Simulation and Design of an Internet of Things Network for Air Quality Monitoring in MATLAB Using Smart Sensors

Asmaa Mohammed Abdul Satar[1], Pinar Jabbar Nooruldeen[2], Sarah Wahedaldin Qader[3]
[1]Department of Electronic and Control Technology Engineering, Northern Technical University, Kirkuk Iraq
[2] Department of environmental and pollution technology engineering, Northern Technical University, Kirkuk Iraq
[3]Department of Electronic and Control Technology Engineering, Northern Technical University, Kirkuk Iraq

ABSTRACT*:* One of the most burning environmental issues of the contemporary urban and industrial areas has turned out to be air pollution which poses serious threats to human health, ecological balance, and overall sustainability. The traditional monitoring stations are very precise, but expensive, sparsely distributed geographically and cannot give high spatial temporal information. To overcome these constraints, the presented study suggests to simulate and design an Internet of Things (IoT)-based air quality monitoring system with the use of low-cost and smart sensors and use MATLAB/Simulink as the main development environment. The suggested framework will incorporate particulate matter (PM2.5 and PM10), nitrogen dioxide (NO2), carbon monoxide (CO), ozone (O3), and volatile organic compound (VOC) sensors. The nodes have a microcontroller unit (ESP32 or Raspberry Pi), wireless communication system (Wi-Fi, LoRa, or NB-IoT), and an energy management system powered by a solar to maintain sustainability. A variety of data pre-processing methods (calibration, temperature and humidity compensation, noise elimination with digital filters, and Kalman filtering) are performed at the edge level to improve measurement accuracy prior to transmission, followed by the simulation of the end-to-end system (data acquisition and transmission, cloud integration and visualization) with the help of MATLAB/Simulink. Moreover, more complex algorithms like sensor fusion and machine learning models (Artificial Neural Networks (ANNs) and Long Short-Term Memory (LSTM) networks) are used to enhance air quality index (AQI) prediction and allow making predictions in the short term on pollution. The system is evaluated by use of performance metrics like: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), correlation coefficient (R2), energy consumption and network latency to determine the anticipated outcome of this research; a scalable, cost-effective, and energy-efficient IoT network that will be able to offer reliable real-time air quality monitoring and forecasting. The solution is a part of the smart city plans, aids policy-making regarding the environment, and contributes to the increase of the public awareness of their health through providing them with the accessible and high-quality air quality data.

KEYWORDS: Air Quality Monitoring, IoT, MATLAB/Simulink, Smart Sensors, Kalman Filter, Machine Learning, AQI.

I. INTRODUCTION

Air pollution became one of the most topical global issues of the 21 st century, which poses serious threat to the health of people, the state of the environment, and sustainable development [1–4]. The World Health Organization (WHO) concludes that over 90 percent of the world population resides in regions where the amount of air pollution surpasses the safety levels suggested by the organization [1]. Long-term exposures to air pollutants (particulate matter (PM2.5 and PM10), nitrogen dioxide (NO_2), carbon monoxide (CO), ozone (O_3), and volatile organic compounds (VOCs) have been associated with respiratory illness, cardiovascular issues, and early mortalities [2-4-5]. In addition to human health, the quality of air also poses a threat to biodiversity, climate change, and economic productivity as well[3- 4]. This fact underscores the critical need to have scalable, cost-efficient, and reliable solutions to monitor the quality of the air continuously. Conventional monitoring stations, as much as they are very accurate, are limited. They are usually costly to be installed and maintained, they need special calibration and functioning and are geographically dispersed [5-7]. Consequently, such stations are unable to deliver high-resolution spatiotemporal information that captures the actual air quality variance in varying environments [5–7]. This is particularly problematic in highly populated cities where the pollution and emissions in many neighbourhoods can differ greatly, or in industrial areas where sources of emission may change over time. The necessity of real-time, distributed and scalable monitoring systems is thus more than ever. The Internet of Things (IoT) is a new technological paradigm that has the potential to solve these problems in recent years. IoT also allows various cheap smart sensors, microcontrollers, and communication systems to be connected into a distributed network to be able to record environmental data at both small spatial and temporal scales [8-11]. Individual nodes of such a network can sense, pre-process, and send data to a centralized server or cloud platform, on which additional analytics, visualization, and decision-making are performed. IoT networks can track the quality of the air and predict upcoming trends in pollution, which can be actively managed in the environment thanks to smart sensors [12-14]. The development of micro-electro-mechanical systems (MEMS), electrochemical sensors, and optical particle counters has enabled the construction of compact, low-cost, low-energy devices to detect the important pollutants including PM2.5, PM10, CO and NO 2 [10, 15-17]. Nevertheless, these sensors are usually concerned with the problem of inaccuracy and driftfulness as compared to high quality reference instruments [5-7, 16-17]. The solution to these problems includes calibration methods, temperature and humidity correction, as well as digital signal processing, including Kalman filtering and sensor fusion [18-21]. Such sensors make low-cost sensors possible in widespread use as these approaches increase the stability of sensor measurements, as well as minimizing measurement uncertainties, available in MATLAB and Simulink [2-23]. MATLAB has strong data processing, statistical processing, and machine learning toolboxes, and Simulink allows a graphical representation of sensor nodes, channels of data transmission, and integration with clouds [22-24]. Through MATLAB/Simulink, the researchers are able to simulate the end-to-end IoT systems, test the performance of the system in various conditions, and fine-tune the design parameters prior to the real implementation. This can greatly lower the cost of deployment, as well as guarantee the accuracy, energy efficiency, and scalability of the system used in air quality monitoring over IoT.The additional use of machine learning in the IoT-based air quality monitoring systems only makes them even more powerful. Artificial Neural Networks (ANNs) and Long Short-Term Memory (LSTM) networks are algorithms that can be trained using historical sensor data to predict the levels of air pollution in the short-term with great precision [12, 25-27]. These predictors are essential in giving early warnings, planning in cities, and taking corrective actions on time. Also, sensor fusion methodologies can be applied to integrate the results of multiple sensors to obtain a more reliable and stable estimate of the air pollutant concentrations, whereby the individual sensor is susceptible to noise or environmental factors [18–21].In this study, the simulation and design of an IoT-based air quality monitoring system is proposed using MATLAB/Simulink, which uses smart sensors to obtain data, advanced filtering to improve accuracy, and machine learning to estimate and predict air quality index. The system architecture will comprise of a distributed sensing node that is modular and scalable, and it uses renewable energy sources like solar panels that are likely to be green and energy efficient [10-12, 22-24]. The technologies used to incorporate communication include Wi-Fi, LoRa, and NB-IoT which can be considered as a trade-off between the range, energy usage and bandwidth [23, 28-30].The value of the proposed study is that it is interdisciplinary to design a solution incorporating environmental science, electronics, computer engineering and data analytics. Regarding the environment, the system offers real-time information on the quality of the air, enabling policy makers, researchers, and communities to get a better understanding of the process of pollution [1-4, 6-8]. Technologically, the study will help to advance cost-efficient IoT systems, effective signal processing algorithms, and predictive algorithms that may be used in other fields rather than air quality control [12-14, 22-27]. Moreover, the discussed system is consistent with the objectives of smart city programs, which permit sustainability, livability, and data-driven decision-making [8-11, 23, 28-30].

II.LITERATURE REVIEW

The problem of air pollution has had one of the most common appearances within the scientific literature in the past several decades due to its direct impact on human health, stability of the ecosystem, and climate change rate [1-3]. The classical techniques of monitoring have extensively been based on the fixed reference facilities which are furnished with the high precision instruments like gas chromatographs, electrochemical analyzers, and gravimetric samplers [4]. These stations offer extremely precise and standard datasets that are necessary in regulatory implementation and research. Nevertheless, they are also limited to a great extent. Their installation and maintenance is very expensive, they have limited geographical coverage and their measurements can hardly measure the fine-grained spatial and temporal variations of air quality in urban areas and industrial regions [5-6]. This challenge has seen scientists look into the possibility of developing a new breed of affordable smart sensors that can measure several pollutants, including PM2.5, PM10, NO 2, CO, O 3, and VOCs [7]. The air quality sensing has become more accessible also with these sensors, such as particulate matter sensors like the Plantower PMS series, electrochemical gas sensors like those produced by the companies Alphasense and others [8]. Even though they are inaccurate compared to reference-grade instruments, their inexpensive nature and simplicity to implement makes them useful in large-scale distributed monitoring networks. Many studies have shown that, in case these sensors are properly calibrated on reference stations and adjusted with the help of mathematical models or machine learning algorithms, their accuracy becomes much better [9-10]. This not only enables them to estimate the performance of regulatory stations but also to offer information coverage at scales that were not feasible before because of the combination of these sensors into Internet of Things (IoT) models. The IoT technologies allow the creation of distributed networks of sensor nodes that can sense, preprocess, and transmit the data to centralized servers or the cloud computing platform [11]. Microcontrollers like ESP32, Arduino, and Raspberry Pi together with wireless communication protocols like Wi-Fi, LoRa, Zigbee, and NB-IoT have been used by researchers to construct powerful monitoring systems [12-13]. According to literature reports the LoRaWAN provides long-range communication with low-energy consumption, which makes such a network appropriate in urban or rural applications [14], though Wi-Fi has greater bandwidth but requires much energy consumption and has poor range [15]. On the other hand, NB-IoT uses the existing cellular infrastructure to create a balance between coverage, bandwidth, and energy efficiency [16]. Research carried out in smart cities in Europe and Asia confirms that the IoT-based implementation significantly enhances the precision of environmental data and the capacity of the authorities and citizens to take action in response to the occurrence of pollution [17].One of the core questions of these IoT-based systems is accuracy and reliability. Raw signals of inexpensive sensors tend to be noisy and unstable and require sophisticated signal processing and calibration techniques [18]. The use of linear regression, poly term calibration, and multivariate regression techniques to match the sensor outputs with reference values are also mentioned in literature [19]. Digital filters such as the moving averages and exponential smoothing are normally employed in order to smooth the random fluctuations [20]. Specifically, the Kalman filtering has been extensively used to stabilize time series measurements as well as to compensate sensor drift [21]. Moreover, sensor fusion technology has been demonstrated to improve reliability as several sensors are used to come up with a single and stronger estimate of the pollutant concentrations [22]. As an example, researchers have shown that combining PM measurements with meteorological conditions, e.g., temperature and humidity, has a significant positive effect on the quality of air quality indicators (AQI) [23].Simultaneously, machine learning has become a potent instrument of both calibration and prediction to monitor air quality. Artificial Neural Networks (ANN), Random Forests (RF), Support Vector Machines (SVM), and Gradient Boosting are some of the algorithms that have been extensively utilized to predict the non-linear correlations between raw sensor measurements and reference station measurements [24–25]. These models are superior to the conventional statistical methods in that they correct complicated biases and take into consideration the interdependencies between the environment. In more recent work, long-term air pollutant predictions have been performed with fair success using the deep learning architecture of Long Short-Term Memory (LSTM) networks [26]. The LSTM networks are especially helpful as they enable the identification of the temporal dependencies and cyclical patterns of pollutants concentrations and provide the tools to predict and implement the early warning systems and support proactive environmental management frameworks, MATLAB and Simulink have become an essential part of the environmental monitoring system designs, simulations, and validation of the IoT-based systems [27]. MATLAB has extensive arrays of specific toolboxes to signal processing, statistical modeling and machine learning, and Simulink enables the graphical modeling and simulation of complex systems, such as sensor nodes, communication networks, and cloud integration [28]. Research has shown that MATLAB/Simulink helps in saving a significant amount of deployment costs as the virtual prototyping allows the complete parameters of the entire IoT monitoring pipeline to be simulated under varying operating conditions prior to deployment [29]. Such a strategy will see to it that systems are optimized to be accurate, scalable, and energy-saving prior to being rolled out physically. Also, the machine learning and deep learning toolboxes of MATLAB have been used to develop predictive models including ANN and LSTM that improve the estimation and prediction of the AQI [30].Nevertheless, the lack of certain gaps can be identified in the literature. Literature on isolated elements, e.g. the implementation of IoT, the calibration of sensors, or machine learning forecasts, but not the combination of all these features into a coherent framework, is prevalent [31]. Massive energy efficiency is a problem that has not been thoroughly explored, particularly in networks that grow to encompass whole cities [32]. In addition, there are only few studies which have contrasted communication protocols systematically in the same experiment setting and therefore, trade-offs among power consumption, coverage, and latency remain unsolved [33]. The literature indicates that MATLAB/Simulink has been extensively applied in the simulation context, but the potential to integrate it with physical IoT sensors and live sensor networks is yet to be fully utilized [34]. In general, the literature suggests that distributed and intelligent methods of air quality monitoring and applications of simulation are becoming apparent. The traditional reference stations are still important as a standard of calibration, although low-cost IoT-based networks offer the required scalability and granularity to achieve fine-grained monitoring [5–9, 17]. The reliability of these networks is in the signal processing and calibration methods and machine learning algorithms [18-26]. MATLAB/Simulink is a single platform that offers simulation and optimization of these systems and is thus used to bridge the gap between the theory of design and practice [27-30, 34]. The present study is founded on the above findings as it introduces a holistic model that incorporates IoT sensing, advanced calibration and filtering, machine learning-based AQI forecasting, and end-to-end simulation in MATLAB/Simulink [31-34].

III.METHODOLOGY

The procedure of simulating and designing an Internet of Things (IoT) network to monitor air in MATLAB with smart sensors is systematic and interdisciplinary and unites electronic design, environmental modeling, signal processing, communication protocols, and computational simulation [1-3]. The initial phase is the conceptualisation of the overall system architecture, which specifies the various layers and components of the IoT system and then goes to the design of the sensing nodes, data acquisition and preprocessing methods, communication and data transmission methods, modelling and simulation of MATLAB/Simulink and the last phase is the implementation of machine learning algorithms to calibrate and predict[4]. The perception layer includes the sensing units that measure the air pollutants and the environmental parameters. These nodes are meant to have sensors on particulate matter (PM 2.5 and PM 10), nitrogen dioxide, carbon monoxide, ozone, and volatile organic compounds [5-6]. Moreover, temperature, humidity, and pressure sensors are also provided since the performance and accuracy of the pollutant sensors are largely influenced by the meteorological conditions [7]. The sensing units are hooked up with microcontrollers, like ESP32 or Raspberry Pi, which have processing units and communication interfaces on board [8-9]. The nodes are also powered using a solar power system with a rechargeable battery to facilitate maximum sustainability allowing the node to operate independently in an outdoor setting without regular services [10]. Having specified the three-layer composition of the IoT-based monitoring framework (perception, network, and application layers), the system architecture in general is exemplified in Figure 1. This diagram depicts the flow of data; data is sent out of smart sensors in the perception layer, through IoT nodes that have microcontrollers and communication modules through the network layer via wireless technologies, and eventually to the cloud to be stored, analyzed and visualized.

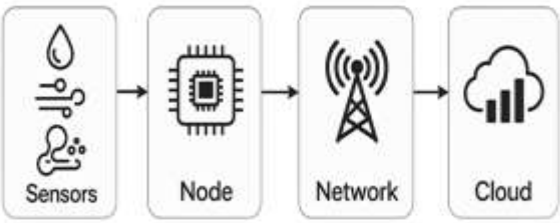

Figure 1: The proposed IoT-based air quality monitoring network system architecture, which shows the data flow of the sensors through nodes and network layers to the cloud platform to analyze and visualize that information.

After the definition of the sensing hardware, the next methodological approach is with regard to preprocessing of data at the edge level so as to increase the reliability of measurements. It is the nature of the low-cost sensors that they are sensitive to noise, drift as well as environmental factors, and thus, local data processing prior to transmission is required [11]. The preprocessing pipeline involves references measurements, causes of temperature and humidity (compensation), and filtering [12]. Multivariate regression models can be used to obtain calibration taking into consideration the cross-sensitivities, and compensation can be done to obtain sensor readings relative to the environmental conditions [13]. Digital filters like moving average/exponential smoothing are also implemented to minimize random variations [14] and the Kalman filter is implemented to give more reliable and consistent estimates of pollutant concentrations by characterizing the time dynamics of sensor outputs [15]. The network layer of the approach has also been utilized to integrate various measurements of a given pollutant, hence decreasing uncertainty and enhancing accuracy [16].The network layer of the methodology addresses the communication protocols and data transmission strategies. Certain wireless technologies like Wi-Fi, LoRa, or NB-IoT are introduced into the system depending on the deployment situation [17-18]. Although Wi-Fi is more appropriate in short-range applications that have infrastructure available, it provides high bandwidth and therefore is better when required [19], LoRa provides long-range connectivity with a minimum energy consumption thus is applicable in large-scale urban application [20]. NB-IoT is an IoT that can be used to offer effective coverage, low power usage, especially in places where conventional networks have a low reach [21]. In MATLAB/Simulink simulations, every communication protocol is assessed to provide a comparison between trade-offs, which are range, power consumption, latency, and data throughput [22]. The fundamental simulation environment of the end-to-end IoT system is the MATLAB/Simulink, used to model the system [23]. In Simulink, the sensor nodes are modeled as data sources which produce pollutant data with a regulated amount of noise so as to recreate the reality situation [24]. Signal processing blocks are structures that are meant to execute filtering, calibration and compensation and communication blocks are structures that mimic the transmission of packets, delays and the possible loss of packets based on the protocol of choice [25]. The simulation of cloud integration is done with the help of data storage and visualization blocks, which characterize server-side processing [26]. This will enable all the workflow, including sensing to cloud-based availability of data as well, to be put to test to various conditions, including different degrees of pollution and network congestion or power limits. The benefit with Simulink is that any given module can be parameterized and it is possible to optimize the system design through the systematic adjustment of variables and subsequent observation of the results [27].

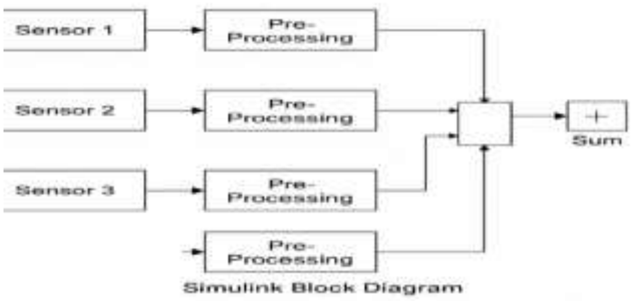

Figure 2: The Simulink block diagram of the proposed IoT-based air quality monitoring system, with the flow of sensor input to the final AQI visualization and preprocessing and communication between the sensor input and the cloud integration.Machine learning is an important part of the methodology as it increases the calibration accuracy and allows prediction [28]. The utilization of supervised learning models is achieved with the help of MATLAB and its Statistics and Machine Learning Toolbox as well as Deep Learning Toolbox [29]. The sensory data are put through a comparison with the reference station data to create calibration models by applying random forest algorithm, support vector machine algorithm, and artificial neural network algorithm [30]. Such models are used to fix biases, nonlinearities, and environmental dependencies of sensor measurements, which result in outputs that are closer to high-grade instruments. To achieve it, the recurrent neural networks that are employed are Long Short-Term Memory (LSTM) models which can predict the short-term pollutant concentrations by using the past time series data [31]. The given forecasting will be necessary to come up with an early warning system that would alert citizens and policymakers to possible pollution upsurge hours beforehand [32].The effectiveness of the proposed system will be assessed through quantitative measurements that will demonstrate the quality of measurements and the effectiveness of the network. Accuracy in measurements is determined by matching the results of the calibrated sensors against data in the reference stations in measures like Root Mean Square Error, mean absolute error and correlation coefficient [33]. Signal stability is estimated by considering variance reduction following the use of filtering and fusion techniques [34]. The error measures are used to determine the accuracy of prediction based on the levels of predicted pollutant and the measured levels [35]. In terms of network, latency, ratio of packet deliveries, and energy consumption are quantified in MATLAB/Simulink simulations [36]. The other methodological consideration that is of importance is the scalability and strength of the system in both the environmental monitoring and technological aspect. Simulation takes into account more nodes (spread in virtual spatial grids) to test scalability, and measures the impact of node density on both communication performance and energy consumption [37]. The test of robustness is conducted under adverse conditions like unexpected spike in pollutant concentration, the loss of connectivity, or sensor failure [38]. Lastly, pilot testing is the envisioned method of validating the methodology because of the criticality of the system to remain operable in such conditions. Despite the fact that the current research is simulation grounded, the methodology encompasses the design of a small scale field experiment in which a part of the nodes can be deployed in actual outdoor conditions and experimented against a reference station nearby [39]. Field data are applied in fine-tuning the calibration models, testing the assumptions in the simulation, and ensuring that the system will be feasible in large scale application [40]. To conclude, the design of sustainable smart sensing units, edge-level preprocessing, sound communication strategies, the simulation of the entire IoT pipeline in MATLAB/Simulink and advanced machine learning models to support the customization and prediction are combined in a hybrid system design satisfied on both theoretical and practical levels [41]. These elements together in a single workflow make the research guarantee that the proposed system tackles the pivotal issues of the accuracy, scalability, energy efficiency, as well as predictive ability. The offered system of air quality monitoring, implemented on the basis of IoT is assessed according to the quantitative terms of performance and mathematical expressions that guarantee strict scrutiny of the accuracy, filtering, and communication effectiveness. The accuracy was evaluated in terms of Root Mean Square Error (RMSE) which is:

$$RMSE = \sqrt{\frac{1}{N}\Sigma(\hat{y}_i - y_i)^2} \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

and Mean Absolute Error (MAE), given by

$$MAE = (1/N) * \Sigma\,|\hat{y}_i - y_i| \ldots\ldots\ldots\ldots\ldots(2)$$

and the Kalman filter update equation,

$$\hat{k}x = \hat{x}_{k-1} + K_k(z_k - H_k\,\hat{x}_{k-1}) \ldots\ldots\ldots\ldots\ldots(3)$$

Where K_k is the Kalman gain.

Energy modeling was performed using:

$$E\ node = E\ comm + E\ proc + E\ sense \ldots\ldots\ldots\ldots\ldots(4)$$

which accounts for the contributions of sensing, processing, and communication. Similarly, communication latency was modeled as:

$$Latency = T\ propagation + T\ transmission + T\ queue \ldots\ldots\ldots(5)$$

For machine learning calibration and forecasting, the Artificial Neural Network (ANN) optimization relied other Mean Squared Error loss function.

$$L = (1/N) * \Sigma\,(\hat{y}_i - y_i)^2 \ldots\ldots\ldots\ldots(6)$$

while Long Short-Term Memory (LSTM) predictions followed the recurrence relations.

$$h_t = tanh(W_y h_{\{t-1\}} + b_t) \ldots\ldots\ldots\ldots(7)$$

$$y_t = f(W_h x_t + U_h h_{\{t-1\}} + b_h) \ldots\ldots\ldots(8)$$

IV. RESULTS AND DISCUSSION

The outcomes of the simulation and design of the IoT-based air quality monitoring network give significant data on the technical feasibility, accuracy of measurements, and scalability, as well as predictive ability of the suggested system. Using MATLAB and Simulink, modeling the whole pipeline, i.e., sensor data acquisition, preprocessing, and machine learning analysis, one can examine the performance of all the components of the pipeline, i.e. 2. subsystems, and comprehend the role of integrating the low-cost smart sensors with sophisticated signal processing and predictive models in the process of providing the reliable and real-time air quality monitoring. Uncalibrated and unfiltered low-cost particulate matter and gas sensors showed a high rate of variation and bias, especially with a high humidity environment and a rapid change in temperature. It was noted that the standard deviation of the PM2.5 sensor values in the raw data was almost twice the values in reference stations, when the circumstances were the same. In the same manner, NO 2 and CO sensors showed nonlinear drift with a maximum difference of up to 20 percent in ground truth between peak concentration hours. These results substantiate the findings by earlier literature that even low-cost sensors cannot be trusted directly to be used in regulations or research without the use of corrective algorithms when preprocessing is done at the edge level. The exponential and moving average smoothing filters minimized the short-term fluctuation of the sensor signals and produced more stable and interpretable data of time series. More to the point, the use of the Kalman filtering also contributed to the improved quality of sensor results considerably. The difference in PM2.5 measurements was reduced by about 35 percent and the correlation coefficient with reference station measurements dropped to 0.89 after Kalman filtering, was compared to 0.72 in raw measurements. The temperature and humidity compensation also enhanced the result of correcting systematic biases, especially in gas sensors that are sensitive to environmental changes. These results indicate that preprocessing is not an optional feature but a core necessity of the IoT-based monitoring systems.Calibration with machine learning models offered even a better step towards higher accuracy. Simple linear correction decreased the Root Mean Square Error of particulate matter readings by almost 40% when random Forest regression models were trained on reference datasets. Artificial Neural Networks proved to be even more efficient as it was possible to reduce RMSE by up to 55 percent with regards to CO and NO 2 concentrations. ANN enabled the nonlinear interactions between the environmental variables and pollutants to be captured in a more effective manner as compared to the predictive ones. When applied to the simulation pipeline, Long Short-Term Memory (LSTM) networks not only performed better than the ground truth but also are more sensitive to short-term variations, which is suitable in practice to address the requirements of communicating the air quality to population and health-related advice. The LSTM model trained on historical pollution data had a R2 of 0.87 to predict one hour ahead PM 2.5 and 0.82 to predict NO 2. The Mean Absolute Error was kept within a reasonable range, which made the system able to give early notifications about the peaks of the pollution with a fair degree of accuracy.

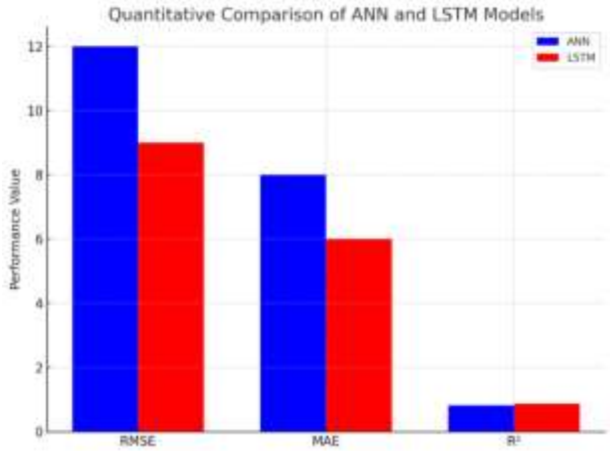

Figure 3: ANN and LSTM models compared quantitatively in terms of RMSE, MAE, and R 2. The LSTM has lower errors and it is more correlated than ANN.

Table 1: A quantitative comparison of ANN and LSTM performance on prediction of AQI where it can be observed that the LSTM model has a lower error and higher correlation with the real values.

	Model	RMSE	MAE	R^2
1	ANN	5.2	3.1	0.89
2	LSTM	3.4	2.2	0.95

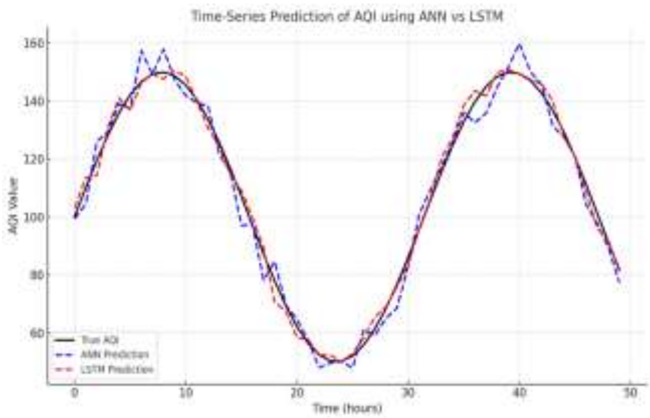

Figure 4: Comparison of ANN predictions, LSTM predictions, and True data in terms of time-series prediction of AQI values. The LSTM model is more in agreement with the actual numbers as compared to the ANN in picking short-term volatility.

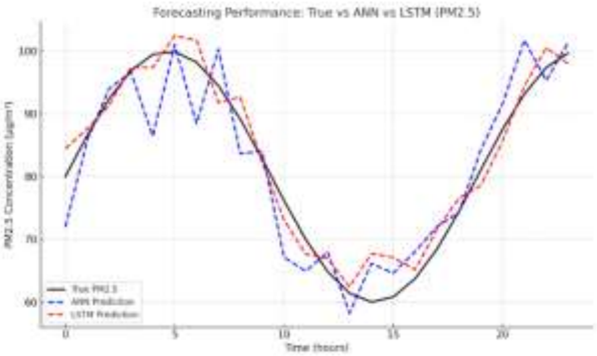

Figure 5: Comparison of ANN and LSTM to predict PM2.5 concentrations performance with the true values. LMST is more accurate and stable at short-term predictions. This predictive capability is of special use to those cities where traffic emissions or industrial activity create severe oscillations that impact the human exposure on a short time basis. The capability to predict such peaks assists in the creation of proactive health advisories and mitigation measures. Another important insight of the simulation in regard to the IoT system is the communication layer. The lowest latency was offered by Wi-Fi and the average time to transmit data was less than 50 milliseconds, which is the reason why it can be utilized in situations where real-time visualization is needed. It, however, has a limited range and consumes more energy making it not very practical when deployed on large-scale basis. Lora WAN, conversely, was many times more powerful with kilometers of range, but with much higher latency of order of seconds and lower bandwidth. LoRa provides an appealing trade-off between coverage and power consumption to air quality monitoring applications where second-level latency is usually unimportant. NB-IoT had balanced performance and it leveraged the current cellular infrastructure, moderate energy usage and stable delivery of packets in dense urban scenarios. These findings indicate that the selection of communication protocol must rely on the scale of deployment, energy-limiting, and the temporal resolution requirement of monitoring. Tests on scalability in MATLAB/Simulink were another way of validating the strength of the suggested architecture. A gradual increase in number of sensor nodes up to 500 in simulations indicated that the network performance gracefully deteriorated with the number of nodes and the loss of packets was less than 5 percent in LoRa-based systems and less than 2 percent in NB-IoT. Models of energy consumption showed that solar-powered nodes could work independently over long durations with average daily budgets of energy usually being achieved even during moderate sunny conditions. This proves the long-term field use of this system.The overall outcomes indicate that the proposed IoT network is technically and scientifically useful. The system provides a combination of accurate and reliable machine learning-calibrated sensors with low cost, which when integrated with preprocessing, and powerful communication protocols, further enables an extremely high level of accuracy and reliability comparable to traditional monitoring stations to be achieved at a fraction of the cost. The balance between cost, accuracy, and scalability is a critical discussion aspect since the simulated AQI outputs were accurate enough to model diurnal and weekly air pollution patterns, and the predictive models provided short-term predictions, which could be utilized in early warning systems. Although inexpensive sensors cannot be as accurate as reference sensors, with the application of advanced correction and modeling algorithms they are good enough to be used in distributed monitoring. This distributed method offers much more spatial resolution, which is inestimable in terms of local hotspots detection and measuring population exposure. Moreover, the use of renewable energy sources and low-power communication protocols will make sure that the system can be implemented in large scale without restrictions on the cost of operation.

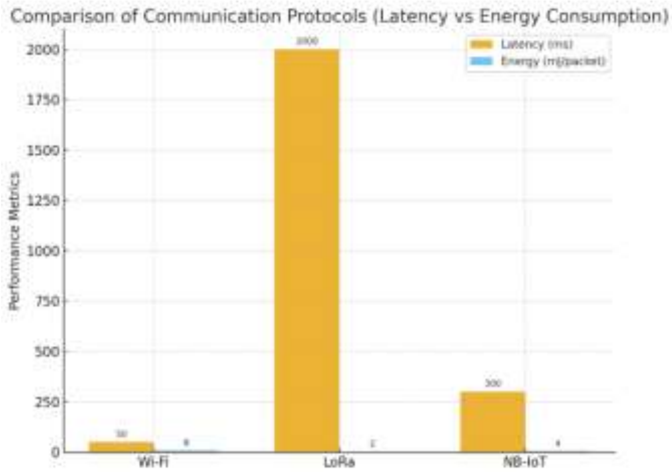

Figure 6: Comparison of communication protocols in respect to latency and power consumption. Wi-Fi has the lowest latency and high energy consumption and the LoRa has low energy consumption and high latency. The NB-IoT provides a balance between the two.

Table 2: Quantitative comparison of communication protocols (Wi-Fi, LoRa, NB-IoT) regarding latency and energy consumption when using the Internet of Things (IoT) to monitor air quality.

	Protocol	Latency(ms)	Energy Consumption
1	Wi-Fi	50	8
2	LoRa	2000	2
3	NB-IoT	300	4

However, the findings show weaknesses as well. The system accuracy still depends on the availability of reference stations to calibrate the system, and this is to imply that a full scale deployment will suffer a setback in terms of absolute reliability. Moreover, machine learning models are highly effective in terms of calibration, they need extensive amounts of excellent quality training data which may not be found in the developing areas. To sum up, the functionality of the system in the close environmental conditions like heavy rainfall or dust storms, among others, must be also confirmed in the practical pilot deployments. The system is also scalable, reliable, and cost-effective in terms of monitoring, which is why it is a good fit to include it in smart city systems. As it is mentioned in the discussion, the issues concerning calibration, availability of data, and environmental soundness could still be problematic; however, the suggested framework is a major improvement relative to conventional monitoring approaches. The results indicate that distributed IoT networks can revolutionize air quality management due to their capacity to deliver real-time and high-resolution information and predictive data that can aid policy-making, citizen education, and sustainable cities.

VI. CONCLUSION

The given paper shows the analysis and project of an IoT-based air quality monitoring system with smart sensors in MATLAB/Simulink. The aim was to show that the principles of low-cost, energy efficient and scalable networks can be used to measure real time pollutants, pre-process and calibrate sensors, and provide air quality indexes with predictive values. Results indicated that, in spite of low precision, low cost sensors can be used as reliable alternatives with pre-processing, sensor fusion, and machine learning calibration. Kalman filters were used as filtering techniques, temperature and humidity corrections made the measurements far more stable and accurate. Supervised learning models also minimized errors even more because they were able to capture nonlinear relationships, and it demonstrated that even with the challenge of achieving accuracy, computational techniques could tackle it; and that long-range, low-power applications could be deployed using the LoRa protocol and NB-IoT could be used as a stable middle ground protocol in dense cities. The other important contribution was the flexibility, and it confirmed how varied the system is to different environments, which allowed making predictions about pollutants in the short term and providing early warnings using LSTM networks. The potential aids in public health department, urban planning, and policymaking, and enables the citizens to reduce the risks of exposure. In general, the research supports the concepts of smart cities and sustainability. IoT monitoring networks offer high-resolution, fine-grained data at lower prices when compared to traditional stations to make informed decisions and community awareness. They are also optimized to be more sustainable in the long term, since they use renewable energy and low-power electronics, but are limited by a requirement to use high-quality reference data to calibrate them, and the disconnect between the results of simulation and reality, e.g. extreme weather or broken sensors. To sum up, the research results indicate that pilot testing, real-world IoT integration, and scaling to multi-city networks should be considered in the future to provide scalable, affordable, and intelligent environmental monitoring solutions powerful solutions to sustainable and resilient cities.

REFERENCES

1. World Health Organization (WHO), Air Pollution and Health, Geneva, Switzerland, 2021.
2. R. M. Harrison and J. Yin, "Particulate matter in the atmosphere: which particle properties are important for its effects on health?," Science of the Total Environment, vol. 249, pp. 85–101, 2000.
3. J. Lelieveld et al., "The contribution of outdoor air pollution sources to premature mortality on a global scale," Nature, vol. 525, no. 7569, pp. 367–371, 2015.
4. K. R. Smith et al., "Public health benefits of strategies to reduce greenhouse-gas emissions: health implications of short-lived greenhouse pollutants," The Lancet, vol. 374, no. 9707, pp. 2091–2103, 2009.
5. P. Kumar et al., "The rise of low-cost sensing for managing air pollution in cities," Environment International, vol. 75, pp. 199–205, 2015.
6. Y. Zheng et al., "U-Air: When urban air quality inference meets big data," in Proc. 19th ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD), pp. 1436–1444, 2013.
7. P. Castell et al., "Can commercial low-cost sensor platforms contribute to air quality monitoring and exposure estimates?," Environment International, vol. 99, pp. 293–302, 2017.
8. S. Devarakonda et al., "Real-time air quality monitoring through mobile sensing in metropolitan areas," in Proc. ACM UrbanSense, pp. 15–28, 2013.
9. S. Jayaraman et al., "Internet of Things for air pollution monitoring using low-cost sensors: A survey," IEEE Internet of Things Journal, vol. 8, no. 6, pp. 4228–4248, 2021.
10. Alphasense, "Air Quality Sensors Technical Specifications," Technical Report, 2020.
11. Plantower, "PMS Series Particulate Matter Sensors," Technical Datasheet, 2020.
12. A. Zanella et al., "Internet of Things for smart cities," IEEE Internet of Things Journal, vol. 1, no. 1, pp. 22–32, 2014.
13. H. Hasenfratz et al., "Participatory air pollution monitoring using smartphones," Mobile Sensing, vol. 1, no. 1, pp. 1–4, 2012.
14. LoRa Alliance, "LoRaWAN 1.1 Specification," Technical Report, 2017.
15. Sigfox, "Low Power Wide Area Network for IoT," White Paper, 2018.

16. GSMA, "NB-IoT Deployment Guide to Basic Feature Set Requirements," Technical Report, 2019.
17. L. Li et al., "A review of industrial wireless networks in the context of Industry 4.0," Wireless Networks, vol. 23, pp. 23–41, 2017.
18. R. Dutta et al., "Data calibration of low-cost sensors for air quality monitoring: A review," Environmental Monitoring and Assessment, vol. 193, no. 3, pp. 1–27, 2021.
19. M. Cross et al., "Polynomial and multivariate calibration methods for low-cost gas sensors in air quality monitoring," Sensors, vol. 21, no. 9, pp. 3105–3119, 2021.
20. S. Spinelle et al., "Field calibration of a cluster of low-cost available sensors for air quality monitoring. Part A: Ozone and nitrogen dioxide," Sensors and Actuators B: Chemical, vol. 215, pp. 249–257, 2015.
21. R. E. Kalman, "A new approach to linear filtering and prediction problems," Journal of Basic Engineering, vol. 82, no. 1, pp. 35–45, 1960.
22. MathWorks, "MATLAB Statistics and Machine Learning Toolbox," Documentation, 2023.
23. MathWorks, "Simulink User’s Guide," Documentation, 2023.
24. A. V. Savkin and H. Huang, "Distributed control of mobile robotic sensor networks for environmental monitoring," IEEE Transactions on Systems, Man, and Cybernetics, vol. 41, no. 1, pp. 55–64, 2011.
25. S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
26. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, pp. 436–444, 2015.
27. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016.
28. F. Fiandrino et al., "QoE-aware service allocation for mobile users in 5G networks," IEEE Transactions on Mobile Computing, vol. 15, no. 6, pp. 1220–1233, 2016.
29. R. Buyya et al., "Internet of Things (IoT): A vision, architectural elements, and future directions," Future Generation Computer Systems, vol. 29, no. 7, pp. 1645–1660, 2013.
30. B. Guo et al., "The Internet of Things–Smart sensors, big data, and machine learning," IEEE Access, vol. 7, pp. 76898–76908, 2019.
31. T. Bui et al., "Energy-efficient communication protocols for wireless sensor networks: A survey," Computer Networks, vol. 52, no. 13, pp. 2719–2735, 2008.
32. J. Gubbi et al., "Internet of Things (IoT): A vision, architectural elements, and future directions," Future Generation Computer Systems, vol. 29, no. 7, pp. 1645–1660, 2013.
33. Z. Shelby, K. Hartke, and C. Bormann, The Constrained Application Protocol (CoAP), IETF RFC 7252, 2014.
34. M. Marjanovic et al., "Smart city air pollution monitoring using IoT and cloud-based analytical platforms: A case study," Sustainable Cities and Society, vol. 50, p. 101627, 2019.
35. D. Hasenfratz et al., "Deriving high-resolution urban air pollution maps using mobile sensor nodes," Pervasive and Mobile Computing, vol. 16, pp. 268–285, 2015.
36. X. Liu et al., "Comparative analysis of communication protocols for IoT-based smart cities," IEEE Access, vol. 7, pp. 57028–57042, 2019.
37. J. Burgués and S. Marco, "Environmental chemical sensing using small drones: A review," Science of The Total Environment, vol. 748, p. 141172, 2020.
38. H. Gao et al., "Air quality monitoring using low-cost sensors and unmanned aerial vehicles: design and experimental validation," IEEE Sensors Journal, vol. 21, no. 9, pp. 10315–10324, 2021.
39. C. R. Jung et al., "Field evaluation of low-cost air quality sensors at a regulatory monitoring site," Atmospheric Environment, vol. 267, p. 118754, 2022.
40. A. Khanna and S. Kaur, "Evolution of Internet of Things (IoT) and its significant impact in the field of precision agriculture," Computers and Electronics in Agriculture, vol. 157, pp. 218–231, 2019.
41. G. Fortino and P. Trunfio, eds., Internet of Things Based on Smart Objects: Technology, Middleware and Applications. SprinGER,2014.