

Article

IoT-Based Air Quality Monitoring System for NO₂ and O₃ Concentrations Using Neural Networks in the City of Ibarra

Michael Negrete-Ramírez , Fabián Cuzme-Rodríguez * , Henry Farinango-Endara  and Carlos Vásquez-Ayala 

Facultad de Ingeniería en Ciencias Aplicadas, Universidad Técnica del Norte,
Av. 17 de Julio 5-21 y General José María Córdova, Ibarra 100105, Ecuador;
minegreter@utn.edu.ec (M.N.-R.); hpfarinango@utn.edu.ec (H.F.-E.); cavasquez@utn.edu.ec (C.V.-A.)
* Correspondence: fgcuzme@utn.edu.ec

Abstract

Air quality monitoring is an essential component of urban environmental management; however, medium-sized cities such as Ibarra often lack updated measurements, limiting evidence-based decision-making. This study presents the design and implementation of an Internet of Things (IoT) system for monitoring nitrogen dioxide (NO₂) and ozone (O₃), pollutants closely associated with vehicular traffic and atmospheric degradation. The system integrates calibrated electrochemical sensors, a data acquisition and wireless transmission module, and a web-based platform for real-time data visualization. To improve signal interpretation, a multilayer artificial neural network was implemented and trained over 71 epochs, achieving a classification accuracy of 97.55% on the independent test dataset for three-class air quality level classification. Field measurements conducted at locations with high vehicular flow reported NO₂ concentrations between 0.00 and 0.03 ppm (approximately 0–57 µg/m³) and O₃ concentrations ranging from 15 to 33 ppb (approximately 30–65 µg/m³), remaining below World Health Organization guideline limits. The results indicate that the proposed solution constitutes a low-cost and scalable tool for preliminary urban air quality assessment and real-time classification, particularly useful for municipal environmental agencies, researchers, urban planners, and public health stakeholders in intermediate cities with limited monitoring infrastructure.

Keywords: air quality monitoring; Internet of Things; low-cost air quality sensors; nitrogen dioxide; ozone; artificial neural networks

1. Introduction

Air pollution constitutes one of the main environmental and public health challenges worldwide. The World Health Organization (WHO) estimates that more than 4.2 million premature deaths annually are attributable to exposure to outdoor air pollutants [1]. Among these pollutants, nitrogen dioxide (NO₂) and tropospheric ozone (O₃) stand out due to their high reactivity and their direct association with vehicular emissions and urban photochemical processes, making them critical indicators for air quality assessment [2,3].

In the Ecuadorian context, several institutions have reported a lack of recent studies on atmospheric pollution and a limited availability of equipment for continuous air quality monitoring [4]. Specifically, the city of Ibarra has not had updated air quality records since the 2012–2015 period, mainly due to the deterioration of monitoring equipment and the absence of new measurement programs [5]. This information gap hinders the early identification of increases in NO₂ and O₃ concentrations, particularly in areas with high vehicular traffic, where health risks are significantly higher [6].



Received: 4 July 2026

Revised: 19 August 2026

Accepted: 21 August 2026

Published: 9 September 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

The city of Ibarra was selected not only because of the absence of updated local air quality records, but also because it represents a typical intermediate Andean city in a developing region, where urban growth, vehicular traffic, and limited environmental monitoring infrastructure coexist. This context provides a relevant case study for evaluating low-cost and scalable monitoring alternatives. Although local meteorological and urban conditions may differ among cities, the proposed architecture is modular and can be adapted to other intermediate urban environments by adjusting calibration procedures, deployment height, sampling strategy, and air quality classification thresholds according to local regulations and reference data availability.

In recent years, Internet of Things (IoT) technologies have emerged as an efficient alternative for the real-time acquisition and transmission of environmental data, driven by their low cost, scalability, and integration capabilities with analytical platforms [7]. In parallel, artificial intelligence models—particularly artificial neural networks (ANNs)—have demonstrated strong performance in the interpretation and classification of signals obtained from environmental sensors, providing more robust and accurate estimations under variable operating conditions [8,9].

In addition to standalone IoT sensing platforms, several studies have explored the use of wireless sensor networks (WSNs) for urban air quality monitoring, emphasizing distributed architectures capable of increasing spatial coverage while reducing deployment and operational costs. For example, a WSN-based prototype for air quality measurements in urban areas was reported in [10], demonstrating the feasibility of real-time environmental data acquisition using networked sensor nodes. While these approaches validate the effectiveness of WSN architectures for environmental monitoring, they are generally limited to data acquisition and transmission, without integrating embedded intelligence for in situ air quality classification or exposure-oriented analysis.

The current state of research shows a growing adoption of low-cost sensing technologies combined with IoT architectures to increase spatial coverage and reduce operational costs in air quality monitoring systems [11,12]. Previous studies have shown that these technologies can support urban air quality assessment and exposure-oriented applications, particularly when dense regulatory monitoring networks are unavailable [13–15]. However, low-cost electrochemical sensors are affected by limitations such as sensitivity to temperature and humidity, cross-sensitivity to other gases, long-term drift, and calibration transferability issues [16–19]. These limitations require calibration strategies, co-location validation, or data-driven correction models to improve measurement reliability [20–23].

Recent technical guidance from the World Meteorological Organization (WMO), developed in collaboration with the United Nations Environment Programme (UNEP), emphasizes that low-cost sensor systems should be understood as complementary tools for air quality applications rather than direct replacements for regulatory-grade monitoring stations. These systems can fill spatial gaps in existing monitoring networks, provide localized information in areas without reference instruments, and support source identification, forecasting, and community-level air quality management, provided that their data are verified and interpreted according to their uncertainty and application context [24].

Related Work and Research Gap

To better contextualize the proposed system, Table 1 summarizes representative studies on low-cost air quality sensing, sensor calibration, IoT-based monitoring architectures, and machine-learning-supported air quality assessment. The comparison highlights that, although previous works have demonstrated the feasibility of low-cost sensors, wireless sensing networks, calibration models, and deep learning approaches, several limitations remain

regarding field calibration, sensor drift, environmental cross-sensitivity, computational requirements, and deployment in intermediate cities with limited monitoring infrastructure.

Table 1. Comparison of related works in IoT-based air quality monitoring systems.

Work	Solution	Innovation	Limitations	Unaddressed Aspects
Morawska et al. [13]	Review of low-cost sensing technologies for air quality monitoring and exposure assessment.	Demonstrates the potential of low-cost sensors for increasing spatial resolution and supporting exposure-oriented studies.	Data quality depends strongly on application context, calibration, maintenance, and environmental conditions.	Limited focus on embedded intelligence and real-time classification at the edge.
Castell et al. [14]	Evaluation of commercial low-cost sensor platforms against reference analyzers.	Assesses the contribution of low-cost platforms to air quality monitoring and exposure estimation.	Measurement reliability is affected by environmental variability and requires careful validation.	Does not provide an edge-AI-based classification architecture adapted to low-resource urban contexts.
Karagulian et al. [15]	Review of low-cost air quality sensors for gases and particulate matter.	Provides a broad performance assessment of commercially available low-cost sensors against reference measurements.	Sensor performance varies substantially depending on pollutant type, environmental conditions, and calibration strategy.	Does not propose an embedded IoT architecture with local air quality classification for intermediate cities.
Maag et al. [16]	Survey of calibration methods for air pollution sensor deployments.	Systematizes calibration approaches, error sources, and network recalibration strategies for low-cost sensors.	Highlights the dependency on co-location data and reference instruments for reliable calibration.	Does not address field implementation of a portable NO ₂ /O ₃ IoT node with web-based visualization.
Zimmerman et al. [17]	Random-Forest-based calibration model for low-cost sensors measuring CO, CO ₂ , NO ₂ , and O ₃ .	Demonstrates that machine learning can significantly improve low-cost sensor performance compared with conventional calibration.	Requires reference-grade co-location data and model-specific calibration procedures.	Focuses on calibration accuracy rather than lightweight embedded classification for autonomous IoT deployment.
Bigi et al. [18]	Field evaluation of NO and NO ₂ low-cost electrochemical sensors using three calibration approaches.	Provides real-world evidence of the behavior of NO ₂ low-cost sensors under medium-term deployment conditions.	Sensor response remains affected by drift, environmental interference, and calibration transferability.	Does not integrate cloud communication, embedded classification, or deployment in cities lacking reference stations.
Felici-Castell et al. [25]	Low-cost sensor network (ANN: multi-dense, CNN, LSTM) for CO ₂ , PM, and NO ₂ .	Prediction with 7.2% error; LSTM showed best performance in urban monitoring.	Sensors exhibit high uncertainty under extreme environmental conditions.	Needs continuous calibration and long-term validation of sensor drift.
Pramanik et al. [26]	Thermo-hygrometer + spatial data + LSTM with temporal attention for AQL.	Air quality estimation with minimal infrastructure and very low cost.	Relies on indirect sensors; accuracy depends heavily on auxiliary data.	No direct pollutant measurement; high uncertainty during peak episodes.
Niresi et al. [27]	Graph Neural Networks (GNN) for calibration and spatial-temporal data fusion.	Significant calibration improvement through distributed sensor nodes.	Requires dense networks and high computational resources for implementation.	Physical limitations of sensors persist; requires frequent recalibration.
Shah and Mishra [28]	ANN model for PM _{2.5} prediction designed for deployment on WSN nodes.	Very high predictive accuracy ($R^2 \approx 0.9986$) using lightweight models.	Based on official station data; not directly tested with low-cost sensors.	Ignores gaseous pollutants and spatial variability in distributed IoT.

Overall, the reviewed literature highlights three key ideas that guide the development of the proposed system. First, low-cost sensors can increase spatial coverage and support exposure-oriented air quality assessment, but their use requires explicit consideration of calibration, drift, environmental interference, and uncertainty. Second, IoT and WSN-based architectures provide an effective mechanism for distributed data acquisition and transmission; however, many implementations remain focused on data collection rather than embedded interpretation at the sensing node. Third, machine learning models can

improve the interpretation of low-cost sensor data, but most reported approaches focus on calibration, pollutant forecasting, dense monitoring networks, or data-rich environments with access to reference stations. As summarized in Table 1, practical implementations that combine low-cost NO₂/O₃ sensing, embedded MLP-based classification, wireless transmission, georeferenced web visualization, and deployment in an intermediate city without a local regulatory monitoring station remain limited. This gap motivates the development of the proposed system as a preliminary, scalable, and cost-effective tool for urban air quality assessment.

In line with the WMO/UNEP technical guidance, low-cost sensor networks are most valuable when they complement existing air quality information sources, increase spatial coverage, and support localized environmental assessment in areas where reference-grade stations are sparse or unavailable [24]. However, their application requires explicit consideration of uncertainty, calibration, environmental interference, and validation protocols. Therefore, the proposed system is positioned as a low-cost, preliminary, and exposure-oriented assessment tool for intermediate cities, rather than as a replacement for regulatory monitoring infrastructure.

Based on this research gap, the main objective of this work is to design and implement a low-cost IoT-based system for monitoring NO₂ and O₃ concentrations in the city of Ibarra. The proposed system integrates calibrated electrochemical sensors, wireless data transmission, a web-based visualization platform, and a multilayer artificial neural network for automatic air quality classification. The results show that the proposed system provides a preliminary, scalable, and cost-effective tool for urban air quality assessment in intermediate cities with limited monitoring infrastructure.

The main contributions of this study can be summarized as follows: (i) the design and implementation of a portable low-cost IoT node for real-time monitoring of NO₂ and O₃ concentrations, together with temperature, relative humidity, and georeferenced location data; (ii) the integration of an embedded MLP-based neural network for local air quality classification on resource-constrained hardware; (iii) the validation of the proposed system through spot co-location measurements against REMMAQ reference stations in Quito and a six-month field deployment in the city of Ibarra; and (iv) the demonstration that the proposed architecture constitutes a scalable and cost-effective alternative for restoring air quality surveillance in intermediate cities with limited monitoring infrastructure.

2. Materials and Methods

This section presents the methodology followed for the development of the proposed air quality monitoring system. A bottom-up approach was adopted, and the system requirements were defined in accordance with the IEEE 29148 standard [29], providing a structured framework for design, implementation, and evaluation.

2.1. System Architecture

The proposed system enables real-time air quality monitoring through the measurement of two critical atmospheric pollutants: nitrogen dioxide (NO₂) and tropospheric ozone (O₃). To this end, an autonomous and portable IoT sensor node is implemented, responsible for data acquisition, local processing, and wireless transmission to a web-based platform. In addition, a neural network model is integrated to perform automatic classification of air pollution levels.

The IoT architecture of the system follows a three-layer model designed to ensure efficient information flow from environmental data acquisition to final visualization. Figure 1 illustrates the overall architecture, which is composed of the device layer, network

and support layer, and application layer, each fulfilling specific functions within the system operation. The main characteristics of each layer are described below.

- **Device layer:** This layer constitutes the physical level of data acquisition. It comprises electrochemical sensors for measuring nitrogen dioxide (NO₂) and ozone (O₃) concentrations, as well as auxiliary temperature and humidity sensors to contextualize environmental conditions. A GPS module is also included to associate each measurement with its geographical location, along with the power supply system that ensures continuous operation of the IoT node.
- **Network and support layer:** This layer combines wireless data transmission through a SIM-based GPRS module with edge computing capabilities. The neural network model is executed locally to perform in situ classification of air pollution levels prior to data transmission, ensuring stable communication and optimizing the information delivered to the application layer.
- **Application layer:** The application layer corresponds to the software environment where the data transmitted from the IoT device are managed, stored, and visualized through a web-based interface.

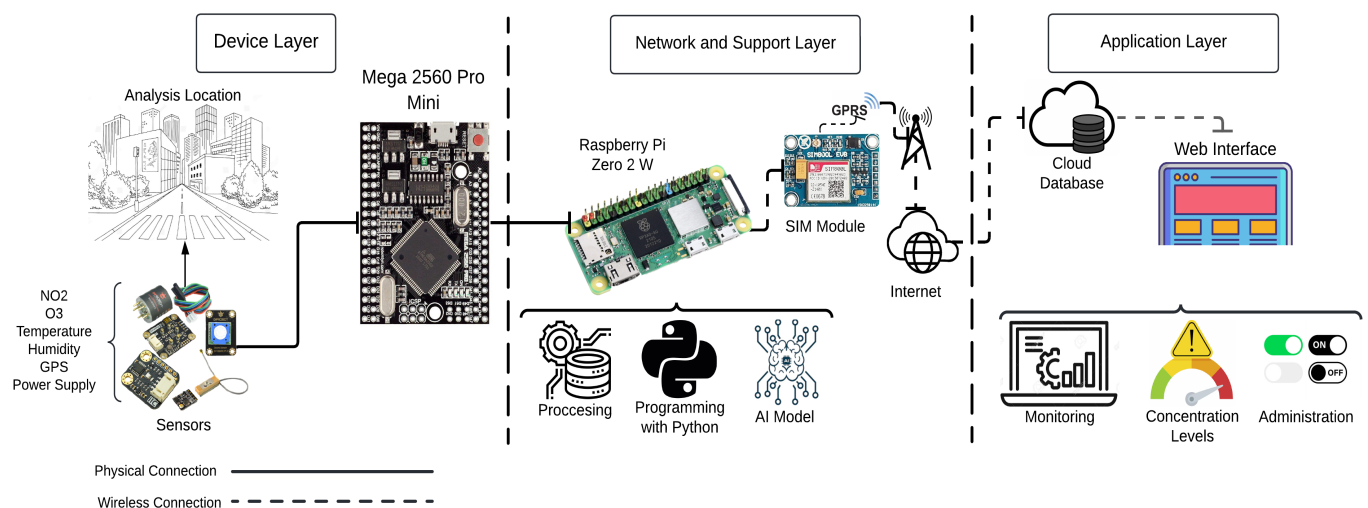


Figure 1. Overall architecture of the proposed IoT-based air quality monitoring system, including the device layer, network and support layer, and application layer. Colors distinguish system modules; solid and dashed lines indicate physical and wireless links, respectively.

2.2. Hardware Design

The proposed monitoring node was implemented using a two-board architecture that separates low-level sensor acquisition from edge processing and wireless communication. This design was adopted to improve modularity, simplify sensor integration, and allow the neural network inference stage to be executed locally before data transmission.

The Mega 2560 Pro Mini ATmega2560-compatible development board (generic module purchased through Amazon Marketplace; manufacturer not specified) was used as the data acquisition unit. This board was selected due to its number of input/output pins, compatibility with I2C, UART, and serial communication interfaces, compact form factor, and ability to operate with 5 V peripheral modules. In the proposed system, the data acquisition board periodically acquires measurements from the NO₂, O₃, temperature, humidity, and GPS modules, structures the sensor readings, and sends the data packet to the Raspberry Pi Zero 2 W (Raspberry Pi Ltd., Cambridge, United Kingdom) through serial communication.

The Raspberry Pi Zero 2 W was used as the local edge-processing unit and IoT gateway. This board executes the preprocessing routines, applies the same normalization parameters obtained during the offline training stage, performs local inference using the trained MLP model, builds the final data packet, and manages the wireless transmission to the cloud platform. This separation between acquisition and processing allows the electrochemical sensors to be managed by a dedicated microcontroller while the neural network and communication routines are executed using Python 3.12.4 on Raspberry Pi OS Lite (release: 21 February 2023; based on Debian 11 “Bullseye”; 32-bit armhf architecture).

The pollutant sensing subsystem consists of two electrochemical gas sensors from the Gravity series. The NO₂ sensor corresponds to the Gravity Factory Calibrated Electrochemical Nitrogen Dioxide Sensor (DFRobot, Shanghai, China), with a nominal measurement range of 0–20 ppm and I2C/UART communication. The O₃ sensor corresponds to the Gravity Electrochemical Ozone Sensor, with a nominal measurement range of 0–10 ppm and I2C communication. According to the manufacturer, both sensors are factory calibrated. However, given the known limitations of low-cost electrochemical sensors, including drift and environmental cross-sensitivity, the present study considers these measurements suitable for preliminary air quality assessment and real-time classification rather than regulatory-grade quantification.

Temperature and relative humidity were measured using a Gravity I2C BME280 environmental sensor (DFRobot, Shanghai, China), while geographic position was obtained using a GPS + BDS BeiDou dual module (Shenzhen Beitian Communication Co., Ltd., Shenzhen, China). These auxiliary variables support microclimatic interpretation and georeferenced visualization but only NO₂, O₃, temperature, and relative humidity were used as input features for the MLP classifier. The GPS coordinates were used exclusively for spatial mapping and data publication.

Wireless communication was implemented through a SIM800L GSM/GPRS module (SIMCom Wireless Solutions Ltd., Shanghai, China). Data were transmitted to the cloud platform using HTTP requests, enabling operation without dependence on local Wi-Fi infrastructure. Each transmitted packet included validated pollutant concentrations, temperature, relative humidity, GPS coordinates, timestamp, and the predicted air quality category.

Case Design

As shown in Figure 2, a three-dimensional (3D) enclosure was designed to house the electronic components and facilitate device portability. The design was developed considering the dimensions of each component as well as the required air intake for the sensors. Dedicated spaces were allocated for each module, including openings for cable routing and connectors.

The main characteristics of the internal design are described as follows:

1. Protective enclosure: This element constitutes the main body and outer shell of the device. Its primary function is to protect the internal electronic components from external factors such as rain and mechanical impacts.
2. Board layer: This layer is designed to accommodate both the data acquisition and processing boards, along with the system power supply. Its purpose is to ensure physical separation between the control electronics and the sensor layer, thereby minimizing interference and organizing the components efficiently.
3. Sensor layer: This structure is dedicated to the mounting and positioning of the air quality sensors, together with the connection and localization modules. Its placement and design aim to thermally isolate the sensors from heat-generating electronic components and to ensure that the sensitive elements of the sensors are directly exposed

to the airflow entering through the ventilation openings, thereby supporting more stable sensor exposure to the incoming airflow.

4. Ventilation openings: These openings serve as the air inlet and outlet between the external environment and the sensor compartment. They are designed with a louvered or zig-zag structure to allow airflow toward the sensor layer (3) while providing a physical barrier against the direct entry of rain, coarse dust, or sunlight, thus preserving system integrity and measurement reliability.

The enclosure was based on an IP55 protective box adapted for portable outdoor operation. Internal 3D-printed supports were used to organize the sensing, processing, communication, and power modules, while the modified external cover allowed environmental air exchange with the sensor compartment. The design aimed to protect the electronics from rain, dust, and mechanical impacts while maintaining passive airflow toward the electrochemical sensors.

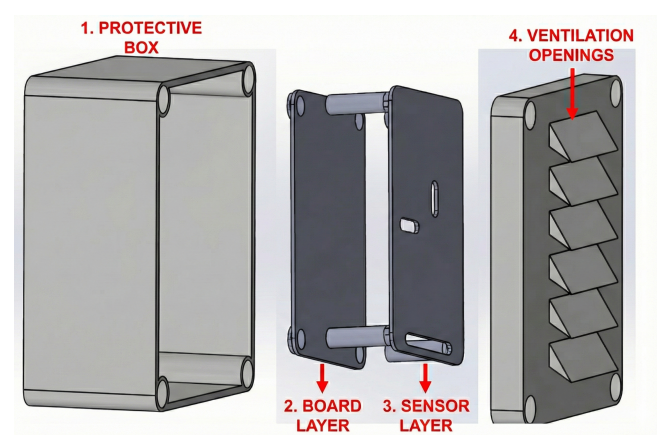


Figure 2. Three-dimensional design of the portable enclosure, showing the protective housing, board layer, sensor layer, and ventilation openings.

Nevertheless, any protective enclosure may influence sensor response by modifying airflow dynamics, solar radiation exposure, and internal temperature. To mitigate this effect, the sensor layer was physically separated from heat-generating electronics, and the gas sensors were positioned close to the louvered ventilation openings. The complete prototype, including the enclosure, was evaluated during component-level tests and field deployment. However, a dedicated controlled experiment comparing sensor response inside and outside the enclosure under standardized airflow and radiation conditions was not performed. Therefore, enclosure-induced airflow and radiative effects are acknowledged as limitations of the current prototype and will be addressed in future stationary versions through improved radiation shielding, calibrated passive or forced ventilation, and extended co-location testing.

2.3. Operational Logic Design

This subsection describes the operational logic governing the interaction among system components, defining the sequence of data acquisition, processing, transmission, and visualization within the proposed IoT-based air quality monitoring system.

2.3.1. Block Diagram

Figure 3 presents the overall block diagram of the system, organized into three main modules that define the complete operational flow: data acquisition, data processing, and data visualization.

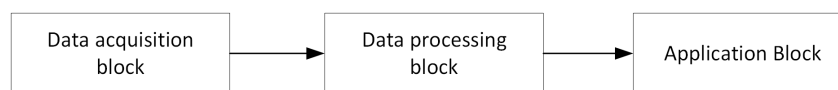


Figure 3. Functional block diagram of the proposed system, including data acquisition, data processing, and application modules.

- The data acquisition block includes the power supply unit and the sensors for NO₂, O₃, temperature, humidity, and GPS. The Mega 2560 Pro Mini development board periodically collects these variables and transmits the structured data to the Raspberry Pi Zero 2 W.
- The processing block is executed on the Raspberry Pi, which performs data preprocessing and applies the trained neural network model to estimate air pollution levels. The processed data are then transmitted to the cloud using HTTP requests via a wireless communication link.
- The application block comprises the web-based interface developed using Flutter, which enables real-time visualization of pollutant concentrations and air quality categories, as well as basic system management functions.

2.3.2. Flowchart

The operational logic of the monitoring system is detailed in the flowchart shown in Figure 4. The process begins with the power-up of the IoT node, followed by a stabilization (warm-up) period of 5 min, which is critical to ensure the accuracy of the electrochemical sensors. Once the thermal and chemical stabilization stage is completed, the system initiates the network connection protocol. If the connection is not successfully established, the device enters a retry loop until connectivity is ensured.

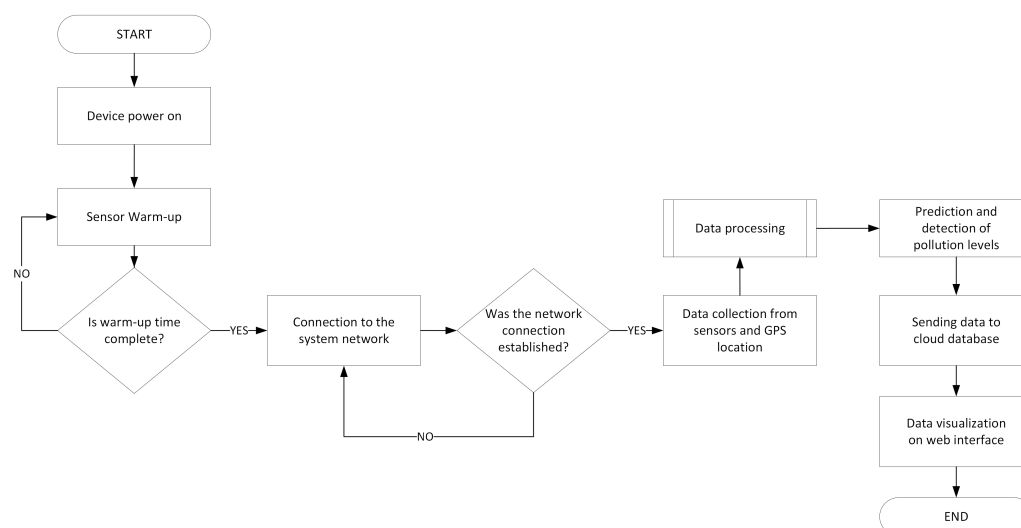


Figure 4. Operational flowchart of the IoT-based air quality monitoring system, from sensor warm-up and network connection to data processing, transmission, and visualization.

With an active connection, the system proceeds to acquire environmental variables from the NO₂, O₃, temperature, relative humidity, and GPS sensors at a sampling rate of one sample per minute. The acquired data are transmitted from the Mega 2560 Pro Mini development board to the Raspberry Pi Zero 2 W, where validation and preprocessing routines are applied. Records with missing values, incomplete packets, or invalid sensor readings are discarded prior to analysis. The four variables used as input to the neural network model are NO₂, O₃, temperature, and relative humidity, while GPS data are used exclusively for georeferencing and spatial visualization. Before inference, the input

variables are normalized using the same StandardScaler parameters obtained during the offline training stage. The normalized feature vector is then processed locally by the embedded neural network model to estimate the corresponding air quality category. Finally, the validated measurements, predicted class, timestamp, and geographic coordinates are transmitted to the cloud server, where the platform manages historical storage and real-time publication through the web interface.

2.3.3. Sequence Diagram

The communication dynamics among the different system modules are illustrated in the sequence diagram shown in Figure 5. The operational cycle starts with node initialization, which triggers data acquisition requests to the environmental sensors (NO₂, O₃, temperature, humidity) and the GPS module. The acquired raw data are transferred to the processing core, where validation and cleaning routines are applied to ensure data consistency prior to analysis.

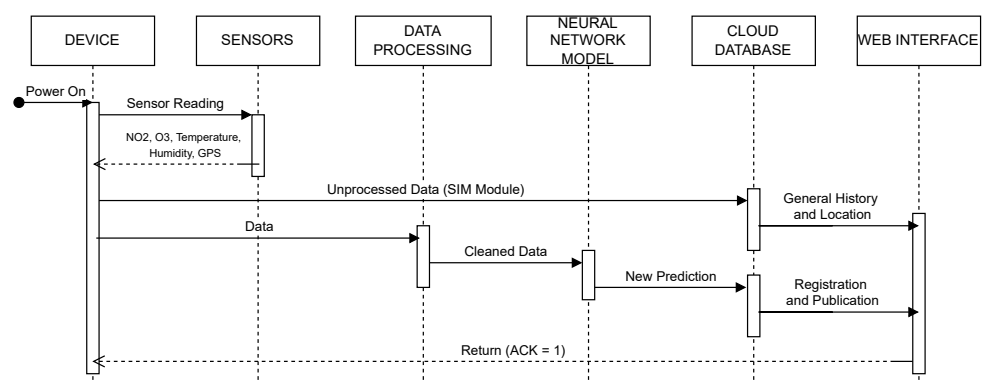


Figure 5. Sequence diagram illustrating the interaction among sensing, processing, cloud storage, and web visualization modules.

Once the data have been processed, the feature vector is input to the local neural network model, which generates an inference regarding the current air quality status. This result, together with the timestamp and geographic location, is transmitted to the cloud infrastructure for database storage and immediate updating of historical and geospatial records.

Finally, the processed information is displayed at the application layer for end-user visualization. The communication cycle concludes with the return of an acknowledgment (ACK) signal from the server to the device. This binary control mechanism validates the integrity of the transaction (value 1), authorizing the acquisition of the next data packet, or, alternatively (value 0), instructing the system to execute a retransmission protocol.

2.4. Neural Network Model Architecture

The predictive core of the proposed system is based on an artificial neural network with a Multilayer Perceptron (MLP) architecture, designed to address a three-class air quality classification problem.

For model training and validation, a balanced dataset consisting of 3270 records (1090 samples per class) was employed. These records were obtained from historical monitoring campaigns, where data entries were pre-classified based on the **Quito Air Quality Index (IQCA)**. This index operates on a “critical pollutant” basis, where the overall air quality category is determined by the pollutant with the highest sub-index. Consequently, the dataset captures the non-linear interactions between NO₂ and O₃, mapping them into three consolidated risk categories:

- **Class 0 (Good):** Conditions compliant with WHO baseline recommendations (IQCA 0–100).
- **Class 1 (Alert):** Moderate pollution levels requiring caution for sensitive groups (IQCA 101–200).
- **Class 2 (Alarm):** High concentrations (e.g., O₃ peaks observed in the dataset) triggering immediate public health warnings (IQCA > 200).

The three-class scheme was adopted as an initial operational classification strategy based on the consolidated IQCA risk categories and the available balanced dataset. This decision avoids over-fragmentation of the training data and prioritizes reliable alert generation under embedded hardware constraints. Although the proposed MLP architecture can be extended to support a larger number of air quality categories, such an extension would require a larger and more diverse labeled dataset, including higher-pollution events, long-term co-location records, and pollutant-specific calibration functions. Therefore, the present version focuses on three consolidated risk levels, while future developments will evaluate finer classification schemes aligned with local and international air quality indices.

Prior to model input, data preprocessing was performed using the StandardScaler technique, normalizing the four input variables to zero mean and unit variance, which is essential to ensure efficient convergence of synaptic weights. The dataset was randomly partitioned using an 80% training and 20% testing split.

The proposed network topology follows a deep sequential structure. The input layer receives four feature vectors corresponding to sensor measurements. Feature extraction is carried out through two fully connected hidden layers, comprising 64 and 32 neurons, respectively. Both hidden layers employ the Rectified Linear Unit (ReLU) activation function due to its computational efficiency and ability to mitigate the vanishing gradient problem. To enhance generalization capability and prevent overfitting, Dropout regularization layers with a rate of 0.2 were applied after each hidden layer, randomly disabling 20% of neurons during each training iteration. The output layer consists of three neurons with a Softmax activation function, producing a probability distribution for multiclass classification.

Model training was configured using the Adam optimizer and the categorical cross-entropy loss function. To improve training efficiency and avoid unnecessary computational cost, two validation-loss-based control strategies were implemented. First, a learning rate reduction strategy (ReduceLROnPlateau) was applied, halving the learning rate after five consecutive epochs without validation loss improvement. Second, an early stopping criterion (EarlyStopping) with a patience of ten epochs was employed, restoring the weights corresponding to the best-performing iteration and ensuring optimal model performance prior to overfitting. The complete configuration of the implemented MLP model is summarized in Table 2.

Table 2. Neural network architecture parameters.

Parameter	Configuration
Architecture	MLP (4-64-32-3)
Total dataset size	3270 samples
Train/Test split	80%/20%
Optimizer	Adam
Loss function	Categorical Cross-Entropy
Activation (Hidden/Output)	ReLU/Softmax
Regularization technique	Dropout (0.2)
Batch size	32
Stopping strategy	EarlyStopping (Patience = 10)

Figure 6 illustrates the sequential data flow of the implemented neural network model. The process begins with the preprocessing of the four input variables using the StandardScaler technique, followed by two fully connected hidden layers composed of 64 and 32 neurons, respectively, both employing ReLU activation functions. Dropout regularization layers with a rate of 0.2 are applied after each hidden layer to mitigate overfitting. Finally, the output layer consists of three neurons with a Softmax activation function, enabling multiclass classification of air quality levels.

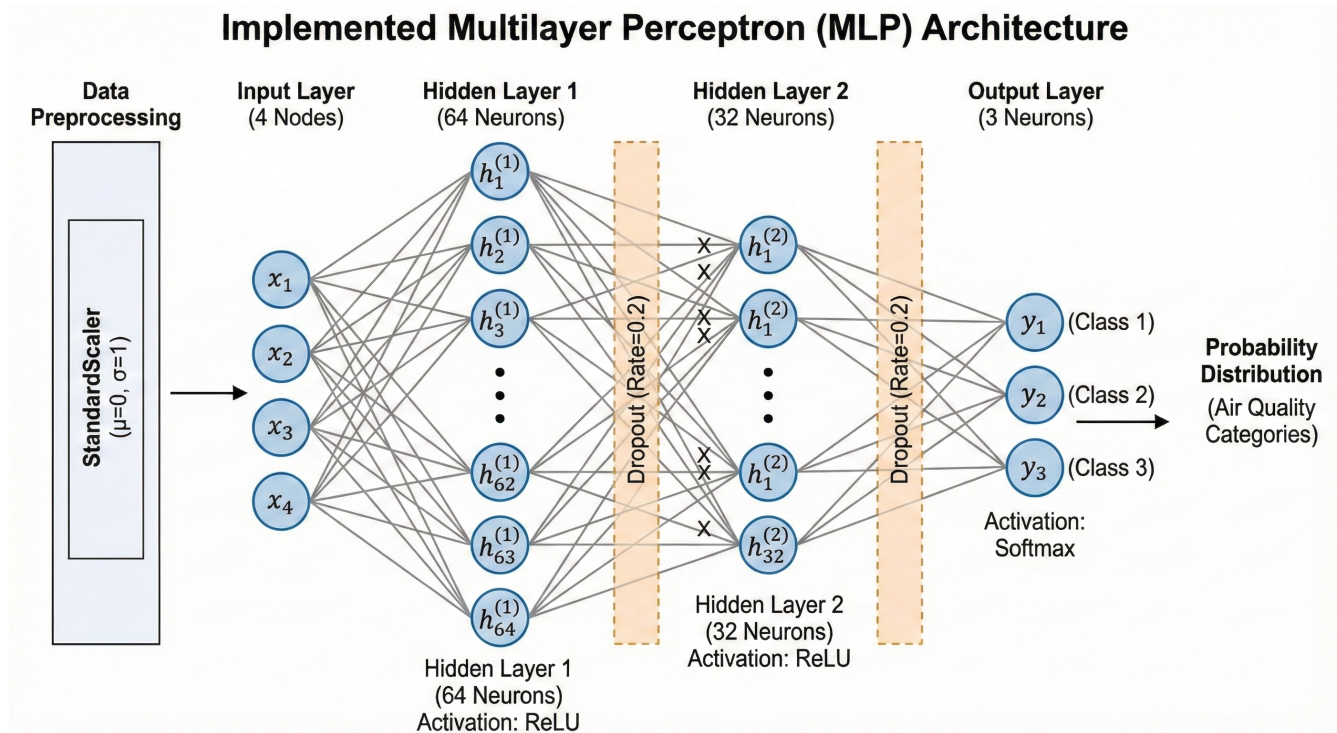


Figure 6. Multilayer perceptron (MLP) architecture used for air quality classification, including input preprocessing, hidden layers, dropout regularization, and softmax output.

2.5. Implementation

Figure 7 presents the electronic design implemented for the air quality monitoring system. The architecture integrates two main control units: an Mega 2560 Pro Mini development board, responsible for the acquisition of environmental sensor data, and a Raspberry Pi Zero 2 W, which performs advanced processing tasks and manages data transmission to the cloud through the SIM800L communication module.

The system power supply is centralized using a portable power bank, which distributes a regulated 5 V output via mini-USB interfaces to the processing units and the SIM800L module. This configuration ensures operational stability and autonomous deployment during monitoring campaigns.

Under the final field deployment configuration, the system operated with a one-sample-per-minute acquisition strategy, which enabled approximately 18 h of autonomous operation during monitoring sessions.

Regarding physical integration, the device enclosure was developed using 3D modeling techniques, achieving a compact and functional internal layout. This spatial arrangement minimizes the overall volume of the prototype while maintaining adequate separation between components and ensuring accessibility to the sensors for future maintenance and calibration tasks, as illustrated in Figure 8.

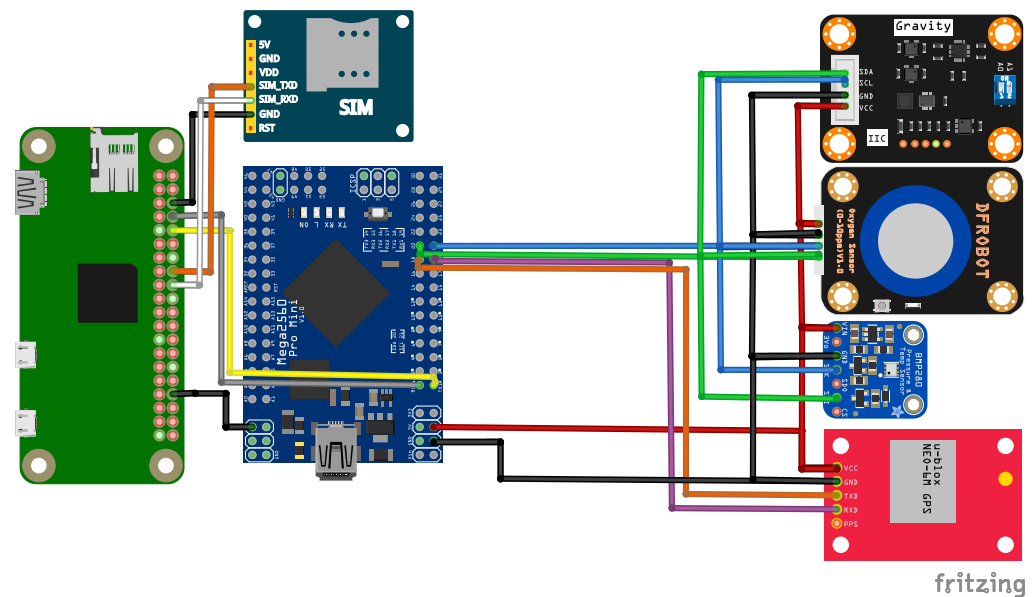


Figure 7. Electronic connection diagram of the implemented monitoring prototype, including sensors, processing units, communication module, and power supply. Diagram generated by the authors using Fritzing.

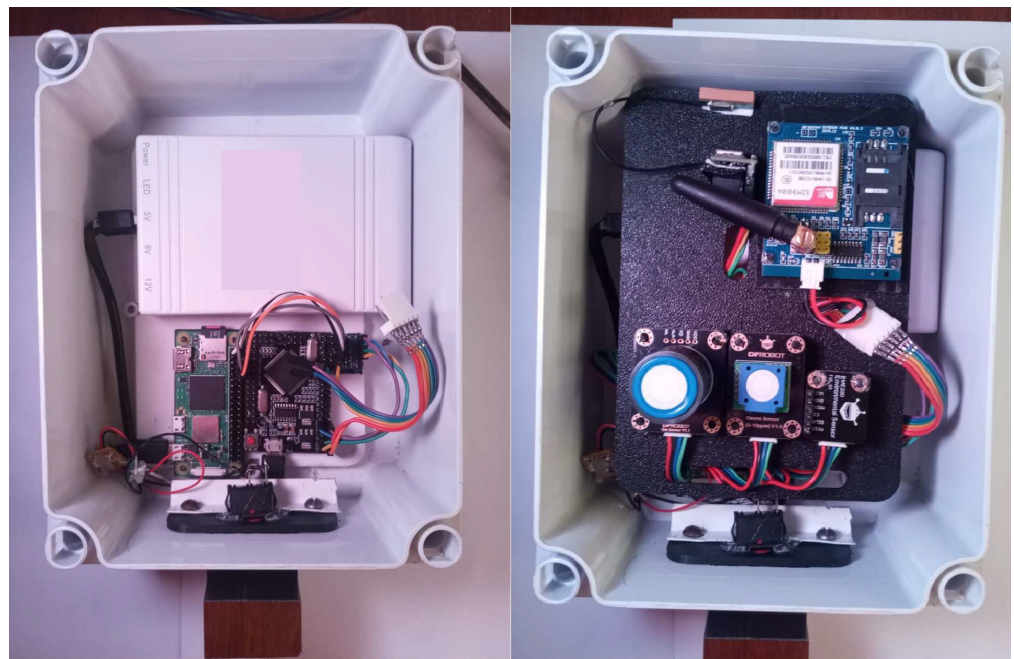


Figure 8. Internal layout of the developed portable monitoring device, showing the arrangement of sensing, processing, communication, and power modules. Photograph taken by the authors.

Figure 9 shows the physical deployment of the implemented prototype at one of the monitoring locations, validating its adaptability and effective integration within the urban environment of the city of Ibarra.

During system operation, the Mega 2560 Pro Mini development board performs periodic acquisition of the environmental variables, while the Raspberry Pi Zero 2 W executes the preprocessing and inference stages locally. Once the validated measurements are normalized, the embedded MLP model generates the corresponding air quality class. The transmitted packet includes the validated pollutant concentrations, meteorological variables, GPS coordinates, timestamp, and predicted air quality category. Data transmission

is performed through the SIM800L communication module, and the acquisition cycle continues only after the reception of a valid acknowledgment signal from the server.



Figure 9. Field deployment of the implemented prototype in the El Arcángel sector, Ibarra, Ecuador, showing the installed sensing node and its outdoor measurement configuration. Photograph taken by the authors.

Although the current prototype prioritizes wireless transmission and web-based visualization, the hardware architecture allows the integration of a low-power LCD, OLED, or e-ink display for local visualization of pollutant concentrations and air quality categories. This option was not implemented in the current version in order to reduce power consumption and maintain the focus on cloud-based monitoring; however, it will be considered in future stationary or semi-stationary versions of the system for field operators and citizen-oriented applications.

2.6. Use of Generative Artificial Intelligence Tools

During the preparation of this manuscript, the authors used ChatGPT (GPT-5.5, OpenAI, San Francisco, CA, USA) for the purposes of English language editing, improvement of scientific clarity, and refinement of editorial style. The tool was not used to generate research data, perform experimental measurements, design the methodology, train or evaluate the neural network model, conduct statistical analyses, create scientific results, or draw conclusions. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

3. Results

3.1. Reference Validation Campaign in Quito

Prior to the long-term deployment in Ibarra, a reference validation campaign was conducted in the city of Quito to assess the response of the proposed IoT air quality monitoring node against the official Metropolitan Air Quality Monitoring Network (REMMAQ). The experimental design considered two complementary scenarios rather than two independent case studies. Quito was used exclusively as a reference-validation scenario because it has regulatory-grade monitoring stations, allowing the low-cost prototype to be compared

against official instruments. In contrast, Ibarra was selected as the target deployment scenario because it lacks updated local air quality monitoring infrastructure. Therefore, the Quito campaign assessed measurement agreement under reference conditions, whereas the Ibarra campaign evaluated operational feasibility, autonomous operation, data transmission, georeferenced visualization, and pollutant behavior in the intended application environment. The validation campaign covered four representative urban locations characterized by different traffic and urban background conditions: Carapungo, Cotocollao, Belisario, and Centro Histórico.

The validation consisted of spot co-location measurements, where the proposed system was operated in close proximity to REMMAQ stations at specific times of the day. Although this approach does not allow the computation of long-term statistical calibration metrics, it provides a practical and realistic assessment of the agreement between the low-cost sensing system and reference-grade instruments under real urban conditions.

Table 3 summarizes the concurrent measurements recorded during the Quito campaign. The results show a consistent agreement in pollutant concentration trends between the proposed IoT system and the official stations. Importantly, the qualitative air quality classification (“Good” or “Alarm”) coincided across all evaluated locations, yielding a 100% categorical match for the spot co-location measurements. This result supports the suitability of the proposed system for preliminary air quality classification and public health-oriented alert applications, while recognizing that extended co-location campaigns are required for quantitative validation and pollutant-specific calibration.

While the categorical agreement was perfect across all locations, some numerical differences in pollutant concentrations were observed. To further quantify these differences, point-wise validation metrics were computed for all cases where both the sensor-based and reference measurements were available. Table 4 reports the corresponding bias and relative error for each pollutant and location.

Table 3. Comparison of in situ measurements between the proposed IoT system and official REMMAQ reference stations during the Quito validation campaign.

Location (Time)	System Source	NO ₂ (µg/m ³)	O ₃ (µg/m ³)	Temp (°C)	Hum (%)	Air Quality Status
Carapungo (08:00)	Designed	21.46	39.19	16.59	55.79	Good
	Official	19.26	19.49	15.96	60.73	Good
Cotocollao (09:00)	Designed	34.38	43.22	21.05	57.70	Alarm
	Official	30.69	–	16.40	68.79	Alarm
Belisario (10:00)	Designed	43.24	54.96	25.40	35.80	Alarm
	Official	34.05	45.21	17.95	49.20	Alarm
Centro (12:00)	Designed	48.19	64.88	28.90	25.20	Alarm
	Official	–	65.86	20.59	34.79	Alarm

Table 4. Point-wise validation metrics between the proposed IoT system and REMMAQ reference stations during the Quito campaign.

Location	Pollutant	Bias (Sensor–Ref)	Relative Error (%)	Unit	Category Match
Carapungo	NO ₂	+2.20	11.4	µg/m ³	Yes
Carapungo	O ₃	+19.70	101.1	µg/m ³	Yes
Cotocollao	NO ₂	+3.69	12.0	µg/m ³	Yes
Belisario	NO ₂	+9.19	27.0	µg/m ³	Yes
Belisario	O ₃	+9.75	21.6	µg/m ³	Yes
Centro	O ₃	−0.98	1.5	µg/m ³	Yes

Due to the spot-based nature of the validation campaign, classical statistical indicators such as RMSE, MAE, or correlation coefficients could not be computed. Instead, the analysis focuses on bias, relative error, and categorical agreement, which are more appropriate for interpreting punctual co-location measurements in a low-cost sensing system designed for preliminary air quality classification and exposure-oriented assessment rather than regulatory-grade quantification.

The relatively high point-wise errors observed in some cases should therefore be interpreted within the limitations of the validation protocol. Spot co-location measurements are sensitive to short-term fluctuations in wind direction, traffic intensity, sampling height, turbulence, and local microenvironmental conditions. In addition, relative error can become disproportionately large when the reference concentration is low, even if the absolute difference remains moderate. These factors explain why numerical deviations may coexist with categorical agreement in the air quality status.

Recent studies on air quality monitoring systems have reported differences ranging from 5 to 15 $\mu\text{g}/\text{m}^3$ when comparing measurements taken at 2 m and 10 m above ground level for traffic-related pollutants [30]. Although these studies focused on particulate matter, the same dispersion principles apply to primary gaseous pollutants such as NO_2 , which tend to accumulate near vehicular emission sources before dispersing vertically. Therefore, the higher concentrations recorded by the proposed system likely reflect more realistic pedestrian-level exposure conditions.

Additionally, discrepancies observed in ambient temperature measurements (e.g., 28.90 °C versus 20.59 °C at the Centro location) are attributed to the direct exposure of the portable device enclosure to solar radiation, in contrast to the controlled radiation-shielded environment of official Stevenson screens. This radiative heating effect may influence the absolute response of low-cost electrochemical sensors and is therefore recognized as a limitation of the current prototype. However, during the Quito validation campaign, the complete device configuration, including the enclosure, maintained categorical agreement with the reference stations. Future work will include controlled enclosure-effect experiments and longer co-location campaigns to quantify the influence of airflow, radiation, and internal temperature on pollutant measurements.

Overall, the Quito validation campaign demonstrates that, despite the absence of long-term co-location data, the proposed IoT system exhibits strong qualitative agreement with reference-grade monitoring stations and provides reliable air quality classification suitable for urban environmental monitoring and public health applications.

It is important to note that the highest relative error was observed for O_3 in Carapungo. This result should be interpreted with caution because the reference concentration was comparatively low; therefore, a moderate absolute difference produced a disproportionately large percentage deviation. In this case, the absolute difference was 19.70 $\mu\text{g}/\text{m}^3$, but the qualitative category remained unchanged. This behavior illustrates one limitation of using relative error as the sole indicator in punctual validation events, especially for low-concentration scenarios. Therefore, the observed discrepancy does not invalidate the practical usefulness of the system for real-time classification, exploratory monitoring, and pedestrian-level exposure assessment. Nevertheless, it highlights the need for longer co-location campaigns to derive pollutant-specific calibration functions and quantify RMSE, MAE, correlation coefficients, uncertainty intervals, and long-term sensor drift.

3.2. Neural Network Model Evaluation

The performance of the proposed neural network architecture was assessed by analyzing the convergence behavior of the loss function and classification accuracy during the training phase. Figure 10 presents the learning curves obtained for both training and

validation datasets. A stable asymptotic convergence can be observed, with closely aligned training and validation curves, indicating the absence of significant overfitting. This behavior is primarily attributed to the effective implementation of Dropout regularization layers and the normalization-based preprocessing of the input data.

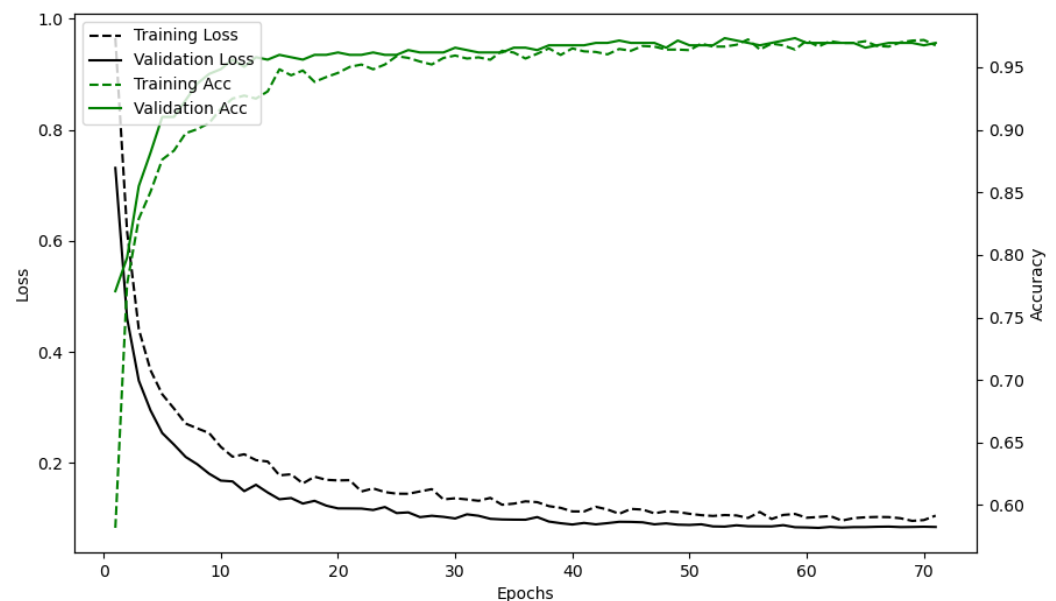


Figure 10. Training and validation learning curves of the proposed neural network model, showing the evolution of loss and classification accuracy over epochs.

The Early Stopping optimization mechanism proved to be effective in managing computational resources. Although the maximum number of epochs was set to 100, the training process automatically terminated at epoch 71 when no further improvement in validation loss was detected, reaching a minimum value of 0.0851. In addition, the ReduceLROnPlateau callback dynamically adjusted the learning rate during the final training stages, enabling fine-tuning of the synaptic weights.

For the final evaluation, the model weights corresponding to epoch 61—identified as the optimal performance point—were restored. When evaluated on the independent test dataset, the model achieved an overall classification accuracy of 97.55% with a loss value of 0.0724. These results indicate a strong generalization capability of the model for classifying air quality levels based on NO₂ and O₃ concentrations, confirming its suitability for deployment on resource-constrained IoT nodes.

The reported accuracy corresponds to 638 correctly classified samples out of 654 test instances. This value is consistent with the class-wise performance metrics presented in Table 5, where the macro-averaged precision, recall, and F1-score reached 0.98. Therefore, the performance of the model is not supported solely by the global accuracy value, but also by balanced class-wise indicators and the distribution of classification errors observed in the confusion matrix.

Table 5. Performance metrics of the MLP-based air quality classification model.

Class	Precision	Recall	F1-Score	Support
Class 0 (Good)	0.97	0.98	0.98	237
Class 1 (Alert)	0.98	0.95	0.96	217
Class 2 (Alarm)	0.98	1.00	0.99	200
Macro Avg.	0.98	0.98	0.98	654
Weighted Avg.	0.98	0.98	0.98	654

Beyond overall accuracy, a more detailed evaluation was conducted using class-wise precision, recall, and F1-score metrics to assess the robustness of the classifier and potential class imbalance effects. Table 5 summarizes these performance indicators for each air quality category.

The results demonstrate balanced classification performance across all air quality categories, with macro-averaged precision, recall, and F1-score values of 0.98. These indicators show that the reported 97.55% accuracy is not the result of a dominant class effect, but reflects consistent performance across the three categories. Notably, the classifier achieved a recall of 1.00 for the highest-risk class, ensuring that all high-risk air pollution events in the test set were correctly identified. This characteristic is particularly relevant for public health-oriented monitoring systems, where false negatives in critical pollution scenarios must be minimized.

Figure 11 presents the confusion matrix obtained from the independent test dataset. The diagonal dominance of the matrix confirms the high consistency of the classifier. Specifically, 232 samples of Class 0, 206 samples of Class 1, and 200 samples of Class 2 were correctly classified, resulting in 638 correct predictions out of 654 test samples. Misclassifications were limited to adjacent categories: five Class 0 samples were classified as Class 1, six Class 1 samples were classified as Class 0, and five Class 1 samples were classified as Class 2. No Class 2 sample was misclassified, which reinforces the suitability of the model for alert-oriented air quality classification.

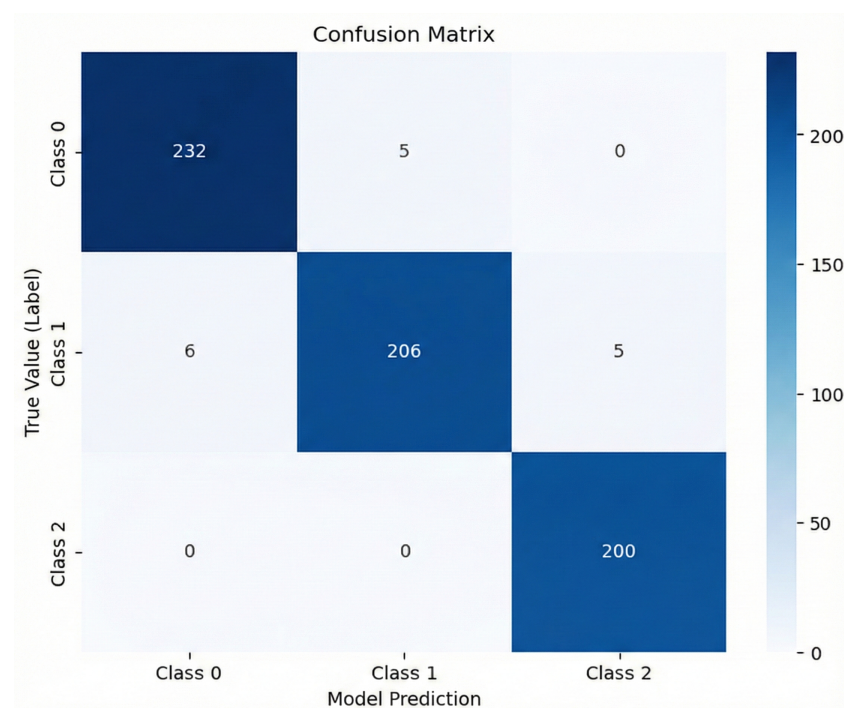


Figure 11. Confusion matrix obtained from the test dataset, illustrating the classification performance of the MLP model across air quality categories.

3.3. Field Tests of the System

Field testing constituted a critical phase for validating the operational performance of the proposed IoT-based NO₂ and O₃ monitoring system under real-world conditions. The prototype was deployed continuously for six months in the city of Ibarra, focusing on areas with high vehicular traffic density, particularly within the urban center, where public and private transportation activity is concentrated according to the local PDOT report.

Data acquisition was conducted following a schedule aligned with periods of peak daily activity. Three time intervals were defined: morning (07:00–10:00), midday

(12:00–14:00), and afternoon (16:00–18:00). These intervals enabled the capture of representative pollutant variations primarily associated with traffic intensity. Throughout the monitoring campaign, the system operated autonomously, collecting environmental data, processing measurements using the embedded neural network, and transmitting classified results to the web platform.

Table 6 summarizes the monthly average values of NO₂, O₃, temperature, and relative humidity recorded between July and December.

Table 6. Monthly average values recorded during field tests.

Month	NO ₂ (µg/m ³)	O ₃ (µg/m ³)	Temperature (°C)	Humidity (%)
July	28.67	39.88	20.30	48.80
August	27.61	53.19	19.68	51.32
September	28.51	52.76	17.44	55.11
October	43.31	54.17	16.44	68.26
November	29.07	45.13	15.95	74.55
December	26.65	59.93	18.26	60.33

Figure 12 illustrates the temporal evolution of NO₂ and O₃ concentrations. NO₂ levels remained relatively stable for most of the monitoring period, with values ranging between 26.7 and 29.1 µg/m³, except in October, when a notable increase to 43.3 µg/m³ was recorded. This increase is likely associated with intensified vehicular traffic and less favorable meteorological conditions for pollutant dispersion.

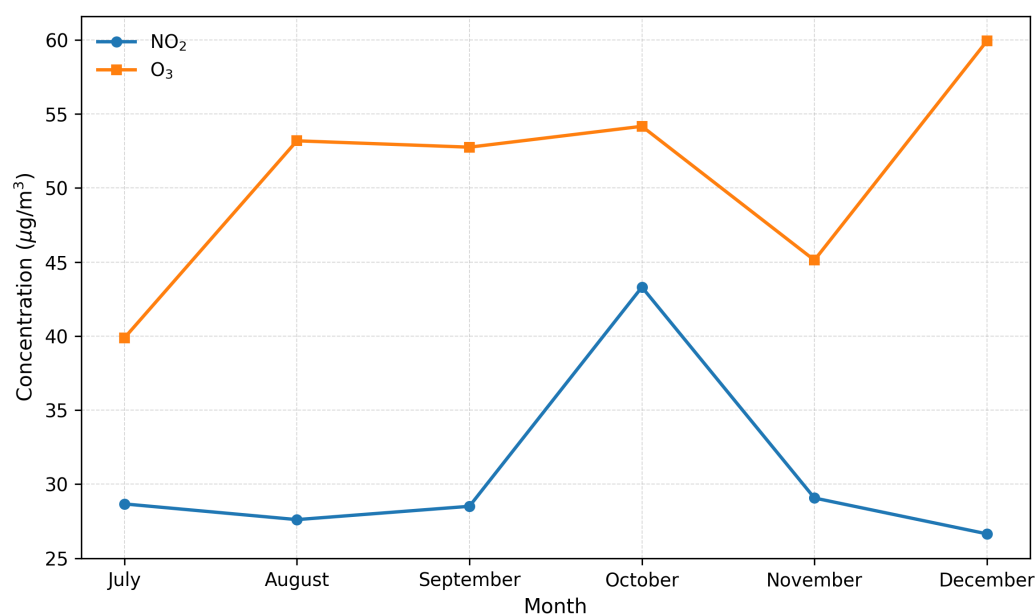


Figure 12. Monthly average concentrations of NO₂ and O₃ recorded during the field monitoring campaign.

In contrast, O₃ concentrations ranged from 39.9 to 59.9 µg/m³, with peak values observed during August, September, October and December. These months coincided with higher solar radiation levels, which promote the photochemical formation of tropospheric ozone.

As shown in Figure 13, ambient temperature exhibited a decreasing trend throughout the monitoring period, declining from 20.3 °C in July to 15.9 °C in November. Conversely, relative humidity increased steadily from 48.8% to 74.6%. These seasonal variations di-

rectly influence the dispersion and chemical reactivity of NO_2 and O_3 , in agreement with established urban atmospheric dynamics.

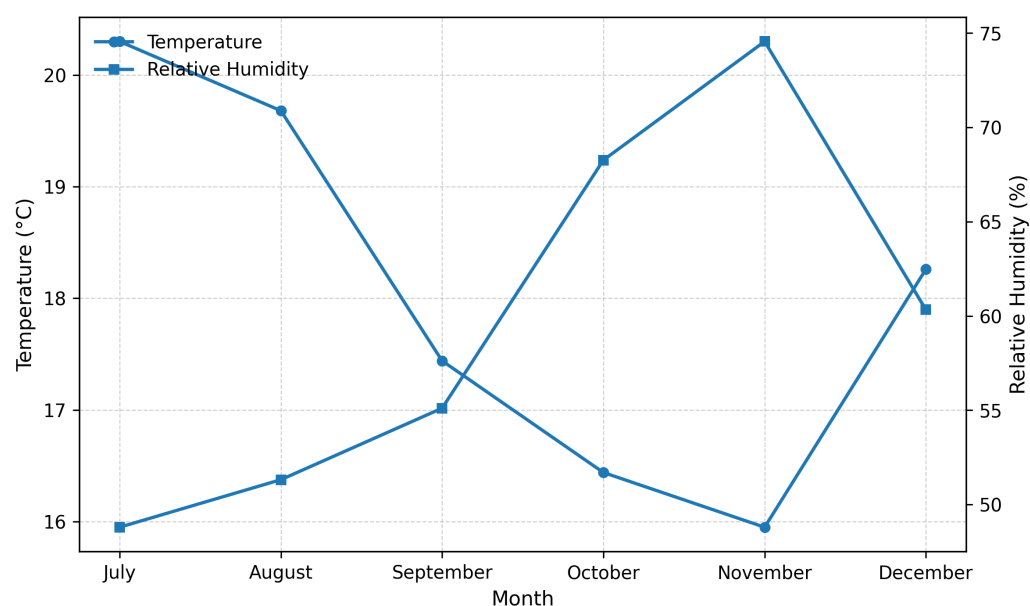


Figure 13. Monthly average temperature and relative humidity measured during the field monitoring campaign.

Figure 14 presents the statistical distribution of monthly NO_2 and O_3 concentrations. O_3 values exhibited a relatively stable median across the evaluated months, with a slight increase observed toward December, reaching an average of $59.9 \mu\text{g}/\text{m}^3$. In contrast, NO_2 displayed higher dispersion, particularly in October, where extreme values of up to $188.1 \mu\text{g}/\text{m}^3$ were recorded. This variability suggests the influence of fluctuating emission sources, likely associated with changes in traffic flow and atmospheric ventilation conditions.

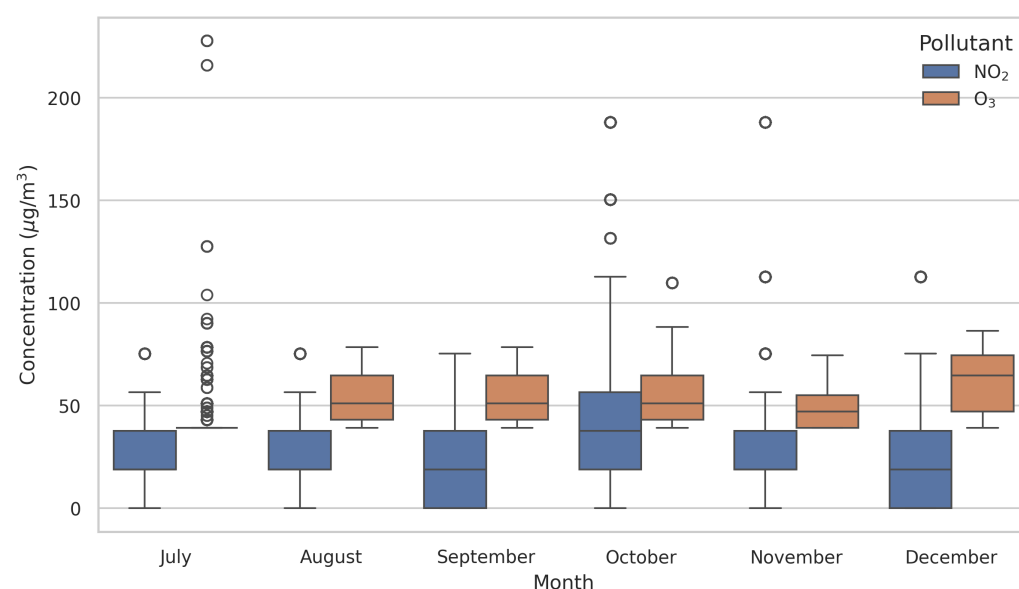


Figure 14. Monthly statistical distribution of recorded NO_2 and O_3 concentrations during the monitoring campaign. Circles represent individual observations identified as potential outliers beyond the whisker range of the boxplots.

The meteorological characterization of the study site is essential to properly interpret the dispersion and formation of atmospheric pollutants. Figure 15 illustrates the monthly

evolution of ambient temperature and relative humidity throughout the monitoring period. A general decreasing trend in average temperature is observed during the central months of the study, with October reporting a mean value of 17.5 °C, coinciding with the highest relative humidity levels, which exceeded 66% during the same month. Toward December, a partial thermal recovery is observed, with maximum temperatures reaching up to 30.5 °C. This meteorological behavior is consistent with the slight reduction in average ozone concentrations observed in November (48.1 $\mu\text{g}/\text{m}^3$), as lower solar radiation and temperature limit the efficiency of photochemical reactions involved in ozone formation.

The analysis of central tendencies and temporal dispersion is summarized in Figure 16, which presents the monthly average concentrations of the monitored pollutants. Tropospheric ozone (O_3) exhibits a gradual increasing trend toward the end of the year, reaching its highest monthly average in December with 59.94 $\mu\text{g}/\text{m}^3$. In contrast, nitrogen dioxide (NO_2) shows its maximum average concentration in October (41.40 $\mu\text{g}/\text{m}^3$), followed by a decrease in December to levels comparable to those observed at the beginning of the monitoring period (approximately 26.6 $\mu\text{g}/\text{m}^3$). The error bars, representing the standard deviation, indicate significantly higher variability in NO_2 concentrations during October ($\sigma = 43.07$), suggesting reduced stability of this pollutant during that month compared to ozone, which exhibited more constrained fluctuations throughout the study period.

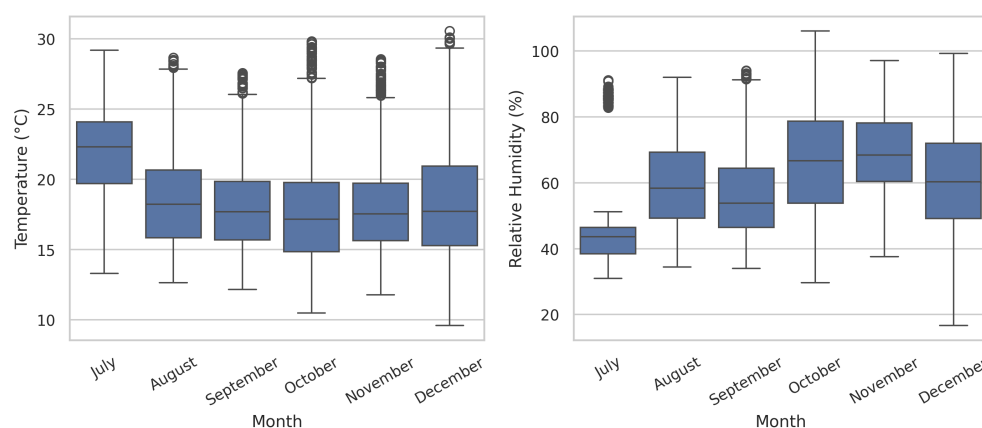


Figure 15. Monthly variation of meteorological parameters, including ambient temperature and relative humidity. Circles represent individual observations identified as potential outliers beyond the whisker range of the boxplots.

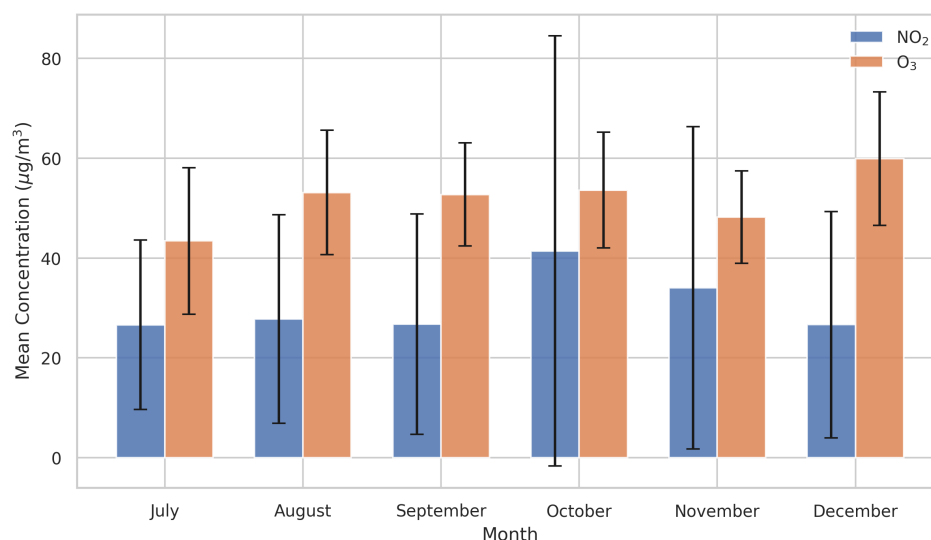


Figure 16. Monthly mean concentrations of NO_2 and O_3 with standard deviation, indicating temporal variability during the monitoring period.

The diurnal behavior of atmospheric pollutants, averaged over the entire monitoring campaign, reflects the typical dynamics of an urban atmosphere. Figure 17 clearly illustrates the inverse relationship between primary and secondary pollutants. Ozone (O_3) follows a unimodal pattern weakly but positively associated with ambient temperature ($r = 0.19$), reaching its maximum average concentration of $59.44 \mu\text{g}/\text{m}^3$ at midday (12:00 h), coinciding with peak solar radiation. Conversely, NO_2 exhibits its highest average concentration of $48.08 \mu\text{g}/\text{m}^3$ during the evening period (18:00 h). This late-afternoon increase in NO_2 is directly linked to vehicular traffic emissions during peak hours, combined with the reduction in the atmospheric mixing layer height at dusk, which favors pollutant accumulation near the surface.

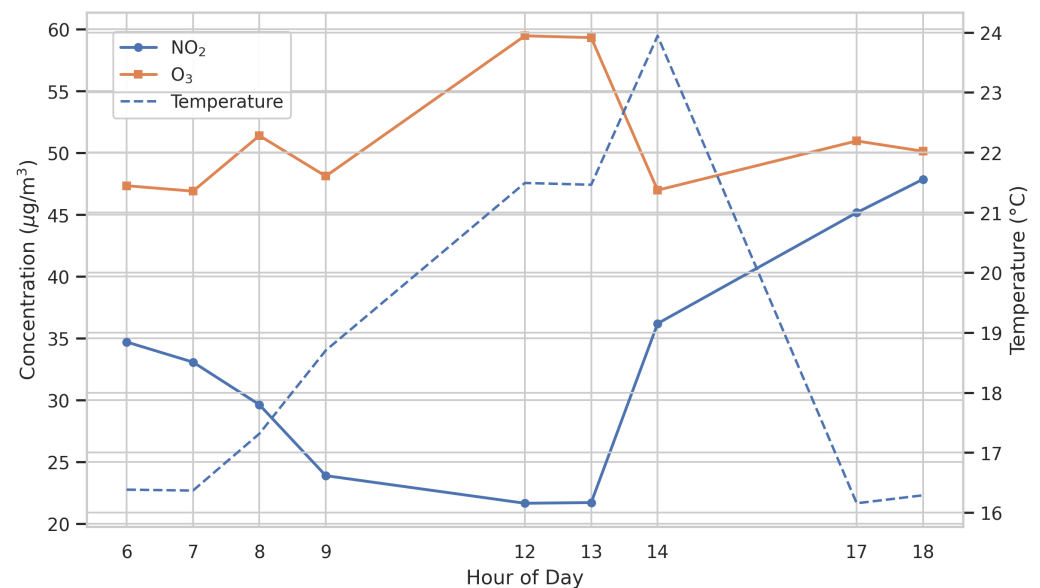


Figure 17. Average daily cycle of NO_2 and O_3 concentrations in relation to ambient temperature.

4. Discussion

The findings of this study support the integration of low-cost IoT technologies with artificial intelligence models as a complementary tool for preliminary air quality assessment in intermediate cities lacking official monitoring infrastructure. The proposed system demonstrated stable operational performance in the monitoring and classification of NO_2 and O_3 , pollutants closely associated with vehicular emissions and photochemical processes, achieving a classification accuracy of 97.55% using a multilayer perceptron (MLP) neural network.

When compared with existing literature, the performance of the proposed system is consistent with prior studies. For instance, the work reported in [25] achieved prediction errors below 8% using low-cost sensors combined with LSTM-based architectures. However, while such approaches emphasize short-term forecasting and typically require high data density, the present study prioritizes real-time classification and operational autonomy. This design choice is particularly suitable for resource-constrained environments such as Ibarra, where system simplicity, robustness, and energy efficiency are critical.

Similarly, although approaches such as AQuaMoHo [26] achieve accurate air quality estimation using LSTM models with attention mechanisms, their reliance on auxiliary external datasets and higher computational demands limits their applicability to low-power embedded devices. In contrast, the MLP architecture implemented in this work provides an effective balance between classification accuracy and computational cost, aligning with findings reported in [21], which highlight the suitability of MLP models

for scenarios characterized by relatively stable environmental patterns and constrained hardware resources.

It should be noted that the present study focused exclusively on the MLP architecture and did not include an exhaustive comparison with classical reference models such as Random Forests or Support Vector Machines. This decision was guided by edge-deployment constraints rather than purely algorithmic performance. Ensemble-based models often require increased memory and processing resources, whereas the MLP offers a compact and optimizable structure suitable for low-cost microcontrollers. Consequently, the validation emphasized operational feasibility within the embedded IoT architecture, under the assumption that the MLP provides an appropriate cost–benefit trade-off for the proposed system.

The purpose of the proposed design is not to outperform reference-grade monitoring stations or more complex forecasting architectures under ideal data-rich conditions. Instead, its contribution lies in providing an integrated, low-cost, portable, and locally deployable solution for intermediate cities where regulatory monitoring infrastructure is absent or limited. Unlike approaches focused exclusively on prediction accuracy, the proposed system integrates direct NO₂/O₃ sensing, embedded MLP-based classification, wireless transmission, georeferenced data publication, and web-based visualization within a single autonomous prototype.

From a cost–benefit perspective, this integration is relevant because it reduces dependence on fixed infrastructure, commercial monitoring equipment, and continuous manual data collection. The use of open-source software tools, low-cost processing boards, electrochemical sensors, and a lightweight MLP model enables practical deployment with moderate technical and economic requirements. Therefore, the value of the system should be interpreted not only in terms of numerical measurement accuracy, but also in terms of accessibility, replicability, operational autonomy, and its capacity to restore preliminary air quality assessment in cities without updated monitoring networks.

From an architectural perspective, the proposed solution aligns with scalability trends described in [12,31], integrating modular hardware, cloud-based services, and web-based visualization. Unlike static monitoring networks or complex vehicular platforms, the portability of the developed sensing node enables rapid redeployment, facilitating exploratory monitoring campaigns in high-traffic areas without reliance on fixed infrastructure.

The portable prototype can be further developed into a stationary or semi-stationary monitoring network by replicating the sensing node at fixed urban locations, particularly near high-traffic roads, schools, transport terminals, and pedestrian exposure zones. In this configuration, each node would preserve the same sensing and edge-classification architecture, while the cloud platform would be extended to manage multiple devices, node identifiers, geospatial dashboards, long-term temporal trends, and device-level diagnostics.

For stationary deployments, several hardware improvements would be required, including solar-assisted or grid-connected power supply, weather-resistant enclosures with improved radiation shielding, calibrated passive or forced ventilation, and periodic co-location calibration against reference instruments. From the software perspective, the current web platform could be expanded to support multi-node data aggregation, spatial interpolation, alert thresholds, maintenance logs, and uncertainty indicators. Therefore, the proposed prototype should be understood as a scalable baseline architecture rather than a closed single-device solution.

Regarding the recorded pollutant concentrations, NO₂ (0.00–0.03 ppm) and O₃ (15–33 ppb) levels remained within the limits established by the World Health Organization, corroborating the historical characterization of Ibarra as a city with relatively low atmospheric pollution. Nevertheless, the absence of a certified reference station for local calibration may introduce bias in absolute concentration values. Despite this limitation, the relative temporal and spatial patterns observed are considered reliable for characterizing environmental variability. As noted in [5,6], the lack of continuous monitoring hinders the detection of emerging trends driven by urban growth, highlighting the relevance of the proposed system as a practical tool for restoring environmental surveillance and supporting data-driven decision-making.

This two-stage validation strategy was necessary because reference-grade stations were not available in Ibarra. The Quito campaign provided a practical benchmark against official instruments, while the Ibarra deployment demonstrated the operational usefulness of the system in the city where updated air quality information is currently unavailable.

In a broader context, this work contributes to narrowing the technological gap identified between regions equipped with advanced monitoring networks and those reliant on low-cost or complementary solutions. This perspective is consistent with recent WMO/UNEP technical guidance, which states that low-cost sensor systems can help fill gaps in existing networks, provide localized insights in areas without reference-grade monitors, and support air quality management when data verification and uncertainty quantification are properly addressed [24]. Although low-cost sensors inherently exhibit limitations related to drift, cross-sensitivity, and environmental interference [19], the integration of local preprocessing and embedded classification can enhance their practical utility for preliminary urban air quality assessment.

Although this study does not develop or assume a climate prediction model, the proposed system is relevant under increasing climatic variability because NO₂ and O₃ dynamics are affected by temperature, relative humidity, solar radiation, and atmospheric dispersion conditions. Therefore, the prototype should be interpreted as a local sensing and classification tool that provides ground-level evidence to complement climate-aware environmental management strategies, rather than as a climate forecasting system. Future multi-node deployments and extended time series will allow the integration of meteorological data, satellite observations, and advanced spatiotemporal models to support more robust analyses under changing climatic conditions.

Finally, while the MLP model demonstrated satisfactory performance within the scope of this study, recent literature suggests that architectures such as CNN–LSTM hybrids or Transformer-based models could further improve analysis in scenarios involving extended time series and multi-node deployments [22,23,32,33]. Future research should therefore focus on expanding the sensor network, integrating advanced spatiotemporal models, and fusing ground-level measurements with satellite observations, following multi-source strategies recommended in [34].

5. Conclusions

The electrochemical sensors used to measure NO₂ and O₃ concentrations exhibited operational stability throughout the field tests, enabling the consistent detection of variations associated with vehicular traffic and ambient conditions. The recorded concentration levels in the city of Ibarra remained within the range classified as “Good” according to World Health Organization guidelines, which is consistent with the characterization of the city as having limited industrial activity and vehicular traffic as the primary emission source.

The implemented artificial neural network significantly enhanced data processing quality by mitigating noise and inherent instabilities typical of electrochemical sensors. Using the best restored weights identified during training, the model achieved a classification accuracy of 97.55%, correctly identifying risk levels associated with NO₂ and O₃ concentrations. These results confirm the suitability of multilayer perceptron (MLP) architectures for low-power IoT applications requiring effective real-time air quality classification.

Field deployments conducted over a six-month period validated the system's performance under real operating conditions, demonstrating sufficient energy autonomy for the monitoring sessions, with approximately 18 h of operation under the one-sample-per-minute acquisition strategy, stable data transmission using the SIM800L communication module, and accurate georeferencing via GPS.

Overall, this study establishes a replicable reference framework for preliminary air quality assessment in intermediate cities with limited monitoring infrastructure. Beyond the technical validation of the prototype, the proposed integration of low-cost electrochemical sensors, embedded MLP-based classification, wireless data transmission, and web-based visualization represents a cost-effective and scalable strategy for supporting evidence-based environmental management. These findings are relevant for municipal environmental agencies, researchers, urban planners, and public health stakeholders seeking complementary tools to increase spatial coverage and restore localized air quality information in resource-constrained urban contexts.

Future work will focus on four main directions. First, extended co-location campaigns will be conducted to derive pollutant-specific calibration functions and quantify RMSE, MAE, correlation coefficients, uncertainty intervals, and long-term sensor drift. Second, the prototype will be expanded into a multi-node stationary or semi-stationary monitoring network to increase spatial coverage across different urban microenvironments. Third, improved enclosure designs, including enhanced radiation shielding and calibrated passive or forced ventilation, will be evaluated to reduce airflow and thermal effects on sensor response. Finally, more advanced spatiotemporal models, including recurrent, hybrid, and graph-based neural network architectures, will be explored when larger datasets and multi-node deployments become available.

Author Contributions: Conceptualization, M.N.-R. and F.C.-R.; methodology, M.N.-R., F.C.-R. and H.F.-E.; software, M.N.-R.; validation, M.N.-R., F.C.-R. and C.V.-A.; formal analysis, M.N.-R. and C.V.-A.; investigation, M.N.-R. and F.C.-R.; resources, F.C.-R. and H.F.-E.; data curation, M.N.-R.; writing—original draft preparation, M.N.-R. and F.C.-R.; writing—review and editing, F.C.-R., H.F.-E. and C.V.-A.; visualization, M.N.-R.; supervision, F.C.-R.; project administration, F.C.-R.; funding acquisition, F.C.-R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Universidad Técnica del Norte, Ecuador, under research grant No. InvestigaUTN-2025-1509.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All source codes developed for this study are publicly available at https://github.com/FirstRSSO/Sistema_IoT_Monitoreo_NO2_O3_UTN (accessed on 2 February 2026).

Acknowledgments: The authors acknowledge the limited use of ChatGPT (GPT-5.5, OpenAI) as an auxiliary tool for English language editing, improvement of scientific clarity, and editorial refinement during manuscript preparation. The authors reviewed and verified all AI-assisted text and take full responsibility for the content of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACK	Acknowledgment
AI	Artificial Intelligence
ANN	Artificial Neural Network
AQI	Air Quality Index
BDS	BeiDou Navigation Satellite System
BME280	Environmental sensor model for pressure, temperature, and humidity
CNN	Convolutional Neural Network
GNN	Graph Neural Network
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile Communications
HTTP	Hypertext Transfer Protocol
IP	Ingress Protection
IQCA	Quito Air Quality Index
IoT	Internet of Things
LCD	Liquid Crystal Display
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MLP	Multilayer Perceptron
NO ₂	Nitrogen Dioxide
O ₃	Ozone
OLED	Organic Light-Emitting Diode
PDOT	Development and Land Use Plan
PM	Particulate Matter
ReLU	Rectified Linear Unit
REMMAQ	Metropolitan Air Quality Monitoring Network
RMSE	Root Mean Square Error
SIM	Subscriber Identity Module
SVM	Support Vector Machine
UNEP	United Nations Environment Programme
WHO	World Health Organization
WMO	World Meteorological Organization
WSN	Wireless Sensor Network

References

1. World Health Organization. What Are the WHO Air Quality Guidelines? 2021. Available online: <https://www.who.int/news-room/feature-stories/detail/what-are-the-who-air-quality-guidelines> (accessed on 2 February 2026).
2. Wang, M.; Li, H.; Huang, S.; Qian, Y.; Steenland, K.; Xie, Y.; Papatheodorou, S.; Shi, L. Short-term Exposure to Nitrogen Dioxide and Mortality: A systematic review and meta-analysis. *Environ. Res.* **2021**, *202*, 111766. [CrossRef] [PubMed]
3. Huang, J.; Li, G.; Xu, G.; Qian, X.; Zhao, Y.; Pan, X.; Huang, J.; Cen, Z.; Liu, Q.; He, T.; et al. The burden of ozone pollution on years of life lost from chronic obstructive pulmonary disease in a city of Yangtze River Delta, China. *Environ. Pollut.* **2018**, *242*, 1266–1273. [CrossRef] [PubMed]
4. Ministerio de Producción, Comercio Exterior, Inversiones y Pesca. *Plan Nacional de la Calidad 2023*; Technical Report; Ministerio de Producción, Comercio Exterior, Inversiones y Pesca: Quito, Ecuador, 2023.
5. GAD Municipal de San Miguel de Ibarra. *Actualización del Plan de Desarrollo y Ordenamiento Territorial del Cantón Ibarra*; Technical Report; GAD Municipal de San Miguel de Ibarra: Ibarra, Ecuador, 2022.
6. American Lung Association. *Nitrogen Dioxide*; American Lung Association: Chicago, IL, USA, 2023.
7. Perera, C.; Barhamgi, M.; Meedeniya, D.A.; Bonilla, V.; Campoverde, B.; Yoo, S.G. A Systematic Literature Review of LoRaWAN: Sensors and Applications. *Sensors* **2023**, *23*, 8440. [CrossRef] [PubMed]
8. Chollet, F. *Deep Learning with Python*, 2nd ed.; Manning Publications: Shelter Island, NY, USA, 2021; pp. 1–504.

9. Hu, Z.; Bai, Z.; Yang, Y.; Zheng, Z.; Bian, K.; Song, L. UAV Aided Aerial-Ground IoT for Air Quality Sensing in Smart City: Architecture, Technologies, and Implementation. *IEEE Netw.* **2019**, *33*, 14–22. [\[CrossRef\]](#)
10. Yamberla-Cacuango, H.; Michilena-Calderón, J.; Farinango-Endara, H.; Cuzme-Rodríguez, F.; Vásquez-Ayala, C. A WSN Network Based Urban Areas Air Quality Measurements Prototype. In *International Conference on Innovation and Research*; Springer: Cham, Switzerland, 2024; pp. 59–69.
11. Yi, W.Y.; Lo, K.M.; Mak, T.; Leung, K.S.; Leung, Y.; Meng, M.L. A Survey of Wireless Sensor Network Based Air Pollution Monitoring Systems. *Sensors* **2015**, *15*, 31392–31427. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Idrees, Z.; Zheng, L. Low cost air pollution monitoring systems: A review of protocols and enabling technologies. *J. Ind. Inf. Integr.* **2020**, *17*, 100123. [\[CrossRef\]](#)
13. Morawska, L.; Thai, P.K.; Liu, X.; Asumadu-Sakyi, A.; Ayoko, G.; Bartonova, A.; Bedini, A.; Chai, F.; Christensen, B.; Dunbabin, M.; et al. Applications of Low-Cost Sensing Technologies for Air Quality Monitoring and Exposure Assessment: How Far Have They Gone? *Environ. Int.* **2018**, *116*, 286–299. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Castell, N.; Dauge, F.R.; Schneider, P.; Vogt, M.; Lerner, U.; Fishbain, B.; Broday, D.; Bartonova, A. Can Commercial Low-Cost Sensor Platforms Contribute to Air Quality Monitoring and Exposure Estimates? *Environ. Int.* **2017**, *99*, 293–302. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Karagulian, F.; Barbieri, M.; Kotsev, A.; Spinelle, L.; Gerboles, M.; Lagler, F.; Redon, N.; Crunaire, S.; Borowiak, A. Review of the Performance of Low-Cost Sensors for Air Quality Monitoring. *Atmosphere* **2019**, *10*, 506. [\[CrossRef\]](#)
16. Maag, B.; Zhou, Z.; Thiele, L. A Survey on Sensor Calibration in Air Pollution Monitoring Deployments. *IEEE Internet Things J.* **2018**, *5*, 4857–4870. [\[CrossRef\]](#)
17. Zimmerman, N.; Presto, A.A.; Kumar, S.P.N.; Gu, J.; Hauryliuk, A.; Robinson, E.S.; Robinson, A.L.; Subramanian, R. A Machine Learning Calibration Model Using Random Forests to Improve Sensor Performance for Lower-Cost Air Quality Monitoring. *Atmos. Meas. Tech.* **2018**, *11*, 291–313. [\[CrossRef\]](#)
18. Bigi, A.; Mueller, M.; Grange, S.K.; Ghermandi, G.; Hueglin, C. Performance of NO, NO₂ Low-Cost Sensors and Three Calibration Approaches within a Real World Application. *Atmos. Meas. Tech.* **2018**, *11*, 3717–3735. [\[CrossRef\]](#)
19. Hayward, I.; Martin, N.A.; Ferracci, V.; Kazemimanesh, M.; Kumar, P. Low-Cost Air Quality Sensors: Biases, Corrections and Challenges in Their Comparability. *Atmosphere* **2024**, *15*, 1523. [\[CrossRef\]](#)
20. Méndez, M.; Merayo, M.G.; Núñez, M. Machine learning algorithms to forecast air quality: A survey. *Artif. Intell. Rev.* **2023**, *56*, 10031–10066. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Mao, Y.; Fang, A.; Chen, Z.; Zheng, Y. Deep Learning for Air Quality Prediction: A Comparative Study. In *2023 5th International Academic Exchange Conference on Science and Technology Innovation (IAECST)*; IEEE: New York, NY, USA, 2023; pp. 1468–1474. [\[CrossRef\]](#)
22. Gilik, A.; Ogrenci, A.S.; Ozmen, A. Air quality prediction using CNN+LSTM-based hybrid deep learning architecture. *Environ. Sci. Pollut. Res.* **2021**, *29*, 11920–11938. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Rath, S.; Maneesha, P. Air Pollution Forecasting using Machine Learning with Temporal Fusion Transformer and Graph Neural Networks. In *3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*; IEEE: New York, NY, USA, 2025; pp. 1798–1803. [\[CrossRef\]](#)
24. World Meteorological Organization; United Nations Environment Programme. *Integrating Low-Cost Sensor Systems and Networks to Enhance Air Quality Applications*; Technical Report; World Meteorological Organization: Geneva, Switzerland, 2024.
25. Felici-Castell, S.; Segura-Garcia, J.; Perez-Solano, J.J.; Fayos-Jordan, R.; Soriano-Asensi, A.; Alcaraz-Calero, J.M. AI-IoT Low-Cost Pollution-Monitoring Sensor Network to Assist Citizens with Respiratory Problems. *Sensors* **2023**, *23*, 9585. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Pramanik, P.; Karmakar, P.; Sharma, P.K.; Chatterjee, S.; Roy, A.; Mandal, S.; Nandi, S.; Chakraborty, S.; Saha, M.; Saha, S. AQuaMoHo: Localized Low-Cost Outdoor Air Quality Sensing over a Thermo-Hygrometer. *ACM Trans. Sens. Netw.* **2022**, *19*, 1–30. [\[CrossRef\]](#)
27. Niresi, K.F.; Zhao, M.; Bissig, H.; Baumann, H.; Fink, O. Spatial-Temporal Graph Attention Fuser for Calibration in IoT Air Pollution Monitoring Systems. In *2023 IEEE Sensors*; IEEE: New York, NY, USA, 2023. [\[CrossRef\]](#)
28. Shah, J.; Mishra, B. Analytical Equations based Prediction Approach for PM_{2.5} using Artificial Neural Network. *SN Appl. Sci.* **2020**, *2*, 1516. [\[CrossRef\]](#)
29. ISO/IEC/IEEE 29148:2018; Systems and Software Engineering—Life Cycle Processes—Requirements Engineering. IEEE: New York, NY, USA, 2018.
30. Murad, J.A.; Ángel Roa Martínez, M.; Martínez, M.R.R. Incidencia de la altura para la medición de la calidad del aire de PM₁₀ y PM_{2.5} como parte del cumplimiento del objetivo y características de los SVCA. *Ing. Compet.* **2023**, *25*, e-21812828. [\[CrossRef\]](#)
31. Alsamrai, O.; Redel-Macias, M.D.; Pinzi, S.; Dorado, M.P. A Systematic Review for Indoor and Outdoor Air Pollution Monitoring Systems Based on Internet of Things. *Sustainability* **2024**, *16*, 4353. [\[CrossRef\]](#)
32. Sharma, E.; Deo, R.C.; Prasad, R.; Parisi, A.V.; Raj, N. Deep Air Quality Forecasts: Suspended Particulate Matter Modeling with Convolutional Neural and Long Short-Term Memory Networks. *IEEE Access* **2020**, *8*, 209503–209516. [\[CrossRef\]](#)

33. Bansal, S.K.; Avula, S.R.; Mehrotra, M.A.; Singh, P.; Gupta, A.K.; Ghosh, P. Machine learning algorithms for predicting air quality index: A case study in urban and industrial zones. *Period. Eng. Nat. Sci.* **2025**, *13*, 425–434. [[CrossRef](#)]
34. Cromar, K.R.; Duncan, B.N.; Bartonova, A.; Benedict, K.; Brauer, M.; Habre, R.; Hagler, G.S.; Haynes, J.A.; Khan, S.; Kilaru, V.; et al. Air Pollution Monitoring for Health Research and Patient Care. An Official American Thoracic Society Workshop Report. *Ann. Am. Thorac. Soc.* **2019**, *16*, 1207–1214. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.