Journal of Agrometeorology
(A publication of Association of Agrometeorologists)
ISSN : 0972-1665 (print), 2583-2980 (online)
Vol. No. 28 (3) : 375-387 (September - 2026)
https://doi.org/10.54386/jam.v28i3.3353
https://journal.agrimetassociation.org/index.php/jam

Research paper

Smart Weather Station for Real-Time Climate Monitoring and Prediction Using IoT and Machine Learning Technologies

JENNY-A. ROSALES-AGREDO[1], FABIÁN-R. JIMÉNEZ-LÓPEZ[1*], and ANDRÉS-F. JIMÉNEZ-LÓPEZ[2]

[1]*Electronic Engineering Program, Faculty of Engineering, Universidad Pedagógica y Tecnológica de Colombia, Research Group I²E, Tunja, Colombia*
[2]*Faculty of Basic Sciences and Engineering, Universidad de los Llanos, Macrypt Research Group, Villavicencio, Colombia*
[*]*Corresponding author email: fabian.jimenez02@uptc.edu.co*

ABSTRACT

Accurate hyper-local climate prediction is essential for decision-making in agriculture, urban planning, and environmental management. This study addresses the scarcity of accessible monitoring solutions by developing an autonomous, IoT-enabled smart weather station that integrates high-frequency data acquisition with real-time Deep Learning forecasting. The hardware architecture utilizes an Arduino™ for multi-sensor data collection—covering temperature, humidity, pressure, wind dynamics, and solar radiation—coupled with LoRa™ for robust wireless transmission. A Raspberry® Pi 3 functions as an edge computing platform for localized processing and inference. Predictive performance was optimized through a multivariate Long Short-Term Memory (LSTM) network designed to capture non-linear temporal dependencies. Operational validation confirmed high data integrity, comparable to high-end commercial stations. The LSTM model significantly outperformed traditional architectures, yielding a Coefficient of Determination (R^2) of 0.957 and a Mean Squared Error (MSE) of 0.0434. This research provides a reliable, low-latency solution for localized forecasting by successfully synthesizing high-precision models for edge deployment. The scalable architecture represents a significant advancement toward actionable, data-driven decision-making in climate-sensitive sectors, offering a low-uncertainty alternative for sustainable resource management and smart initiatives in resource-constrained environments.

Keywords: *Climate prediction, Data processing, Deep learning, Internet of Things, Meteorological sensors, Weather station.*

The variability and dynamics of climatic variables directly influence critical environmental, agricultural, urban, and socioeconomic processes. Precise, real-time meteorological monitoring with high spatial resolution is crucial for understanding microclimatic patterns and supporting informed decision-making in strategic sectors, including agriculture, sustainable urban planning, tourism, and disaster management (Eleftheriou et al., 2018; Huang et al., 2020; Islam et al., 2024; Plein et al., 2025)and this is true around the globe, not only in anyone nation, which might have a devastating effect on agricultural output. Air, precipitation, temperature, humidity, and a host of other environmental variables are subject to constant and unpredictable change. Having a local weather station that can update farmers on the current weather conditions in real-time is absolutely vital. That way, people can protect their crops by making the right decisions at the right time. The quantity of dangerous compounds in air particles has been reduced because of advancements on the Internet of Things (loT. In regions vulnerable to climate change, such as Colombia, the scarcity of reliable, hyper-local climate data introduces significant uncertainty into productive and security activities.

Traditional measurement methodologies, such as Fixed Meteorological Stations (FMS) and manual logging systems, face inherent limitations regarding spatial and temporal resolution, rigidity, high acquisition and maintenance costs, and limited integration with modern real-time data analysis. These constraints severely restrict FMS deployment in resource-constrained environments, necessitating the adoption of emergent technological solutions. The Internet of Things (IoT) has revolutionized environmental monitoring, enabling the development of low-cost, autonomous, and portable smart weather stations (Alam et al., 2025;

Article info - DOI: *https://doi.org/10.54386/jam.v28i3.3353*
Received:29 January 2026; Accepted: 03 August 2026; Published online : 7 September 2026

Hilda et al., 2025; Nasrat, 2025).

These IoT systems combine miniaturized sensors, energy-efficient microcontrollers/SoCs, and wireless connectivity (e.g., WiFi™, LoRa™, cellular) to capture, transmit, and store real-time meteorological variables—including temperature, humidity, precipitation, and wind dynamics (Islam et al., 2024; Jolly et al., 2024; Ngebani et al., 2018)and this is true around the globe, not only in anyone nation, which might have a devastating effect on agricultural output. Air, precipitation, temperature, humidity, and a host of other environmental variables are subject to constant and unpredictable change. Having a local weather station that can update farmers on the current weather conditions in real-time is absolutely vital. That way, people can protect their crops by making the right decisions at the right time. The quantity of dangerous compounds in air particles has been reduced because of advancements on the Internet of Things (loT. This architecture delivers the fine spatial and temporal data granularity necessary for comprehensive microclimate studies (Alves et al., 2024; Azevedo et al., 2025; Budjač & Gašpar, 2024)and its efficient management, based on water balance and meteorological data, requires data collection and transmission systems with high reliability and the capacity to accurately represent the physical phenomena occurring in the agricultural environment. Thus, aiming to manage irrigation using smart devices, this article presents the development of the hardware and software of a complete automatic meteorological station based on IoT, from the calibration of the sensor elements, to the development of the final device, which operates with the routine of reading variables at intervals of 10 s, followed by online storage of average and extreme data every 10 min, followed by estimation of daily evapotranspiration using the Penman–Monteith method. The device has a user communication interface via a messaging application, through which the current weather condition can be requested remotely and receive daily data collected. The final version had its operation analyzed continuously, for a period of one year, together with a commercial automatic station, where a correlation was found between its estimates (R2.

Furthermore, the convergence of IoT and Artificial Intelligence (AI), specifically Machine Learning (ML) and Deep Learning (DL), has significantly improved the processing and prediction of climatic variables compared to traditional statistical methods (Bhattacharjee et al., 2025; Zhang et al., 2025). Modular systems based on microcontrollers (e.g., Arduino™ or ESP32®) and Systems-on-Chip (SoC, e.g., Raspberry Pi®) are widely validated for data acquisition and environmental monitoring (Asanza et al., 2021; Herdianzenda et al., 2021; Hussein et al., 2020; Megantoro et al., 2021). Among the various AI applications for climate time series analysis, Artificial Neural Networks (ANNs) are preferred for capturing nonlinear dependencies, thus enhancing predictive accuracy (Bhatia et al., 2024; Gao, 2025).

Long Short-Term Memory (LSTM) networks are particularly effective, consistently outperforming traditional models (e.g., ARIMA – AutoRegressive Integrated Moving Average) in both univariate and multivariate climate prediction by successfully managing temporal and nonlinear dependencies in meteorological data (Faid et al., 2022; Omar et al., 2021; Perera et al., 2025; Sakthipriya et al., 2024; Wicaksana et al., 2022). Several AI techniques, including hybrid neural network models, have been successfully applied for comprehensive multivariate prediction (Grekhnev et al., 2023; Ioannou et al., 2021; Kalimuthu et al., 2020; Roy, 2020). However, a significant knowledge gap persists: while many existing IoT weather stations successfully focus on data acquisition, there is a critical lack of local, portable, and open-source solutions that cohesively integrate real-time data collection with high-precision AI-based prediction models for generating location-specific forecasts (Hachimi et al., 2023; Luo et al., 2023; Math & Dharwadkar, 2018).

Most current studies focus on analyzing historical data or large-scale prediction, failing to provide a comprehensive, portable solution that combines accessibility with the precision of ML models for local, resource-constrained environments. This study addresses the identified gap by proposing the design and implementation of an intelligent, autonomous, IoT-enabled weather station that cohesively integrates high-precision sensors, local processing, and LSTM-based DL algorithms for the real-time monitoring and prediction of climatic variables. The core innovation is the development of a portable and scalable solution that embeds data acquisition, local processing, and prediction capabilities within a single unit, thereby enhancing accessibility and reliability in remote areas.

The system utilizes low-power hardware platforms and incorporates sensors for temperature, humidity, pressure, solar radiation, and wind dynamics, coupled with wireless communication for autonomy. The ML models are trained on historical and collected data to generate accurate, timely forecasts, facilitating improved decision-making across agricultural, urban, and risk management sectors. The proposed solution was rigorously validated under real-world conditions in a rural environment, with its performance evaluated across various climatic scenarios and benchmarked against conventional and satellite meteorological stations. The implementation of this portable and scalable station was established, and its performance was verified through: 1) Identification of key design features via an exhaustive literature review; 2) Design and integration of a compact hardware/software prototype; 3) Development of ML models trained with relevant historical data; and 4) Validation under real field conditions, contrasting its precision and autonomy against conventional stations and satellite data.

MATERIALS AND METHODS

System Architecture and Hardware Design

The research involved the design and implementation of an intelligent and autonomous meteorological station, leveraging IoT technologies for real-time climate monitoring and prediction. The system architecture was structured as a hybrid, modular model comprising three distinct layers: sensing, transmission, and centralized processing/analysis. This layered approach ensures robust separation of data acquisition and computational tasks, optimizing efficiency and scalability for deployment in remote or resource-constrained environments (See Fig. 1).

The system utilizes embedded hardware, where the

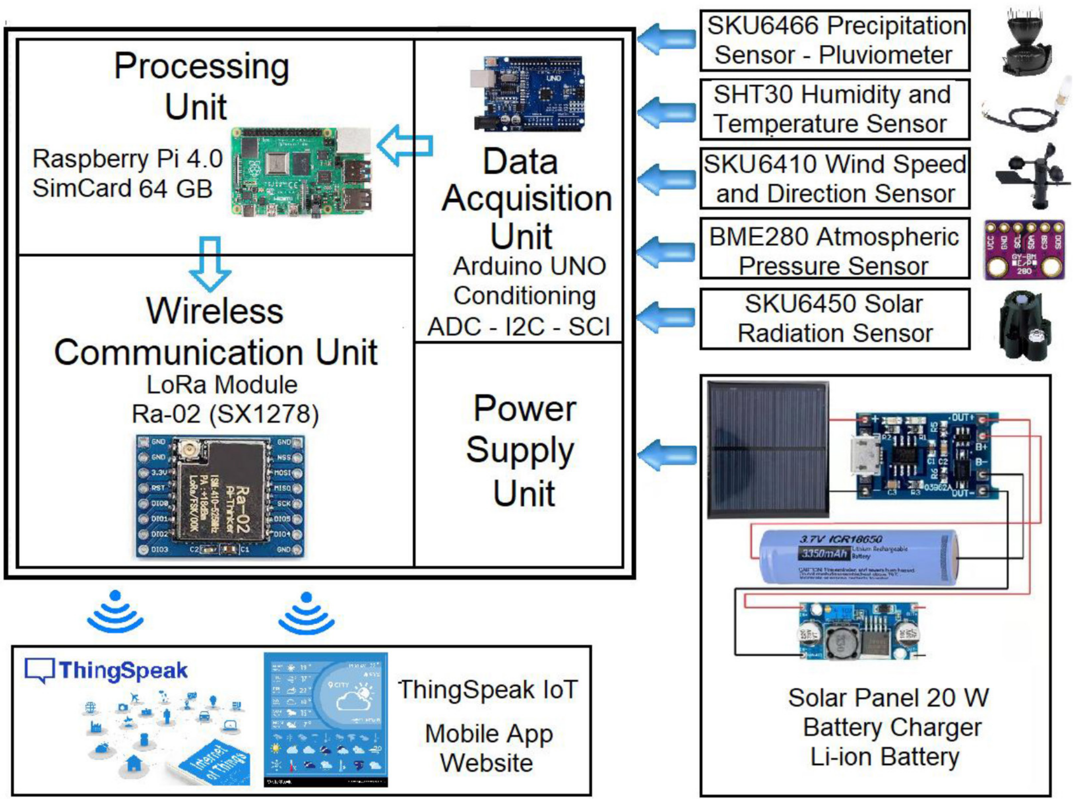

Fig. 1: Smart Weather Station Architecture

primary data acquisition unit is an Arduino™ UNO microcontroller (ATmega328P). This unit is responsible for sensor management and initial data handling. The subsequent tasks of advanced data analysis, storage, and predictive modeling are handled by the core processing unit, a Raspberry® Pi 3 single-board computer (ARM SoC), which functions as the local server.

Wireless communication between the acquisition and processing units is established via LoRa™ Ra-02 modules (SX1278). These modules were selected for their capability to enable reliable, long-range data transmission within the 137 to 525 MHz frequency range while maintaining minimal power consumption. For autonomous operation, the entire station is powered by a rechargeable 12 V Li-ion battery, which is continuously charged by a 20W solar panel system that includes the necessary charging circuit and voltage regulators (3.3 V and 5 V). The electronic components were mounted on a custom-designed 8 cm x 15 cm Printed Circuit Board (PCB), which was designed using LiveWire and PCB Wizard software.

The PCB was coated with a dielectric varnish for protection against moisture and corrosion and secured within a reinforced plastic enclosure finished with white anti-corrosion enamel to minimize solar heating. The external sensors and enclosure were fixed to a durable 2-meter-tall support structure constructed from anodized aluminum and stainless steel, ensuring resistance to weathering and structural stability.

Sensor Instrumentation and Calibration

The station integrates a suite of high-precision sensors, which are connected to the Arduino™ UNO via a combination of analog and digital interfaces, operating at a real-time sampling frequency of once every 10 seconds. To establish rigorous metrological traceability and minimize systematic errors throughout the six-month experimental period, the prototype's sensor suite (operating at a 10 second real-time sampling interval) was periodically cross-validated in a controlled laboratory environment against a certified reference instrument.

A Davis Vantage Pro 2 automated weather station, operated and maintained at Universidad Pedagógica y Tecnológica de Colombia (UPTC), served as the primary benchmark due to its verified accuracy and compliance with international meteorological standards. This systematic cross-comparison framework ensured high calibration reliability, strict data traceability, and strong confidence in the measurement precision of the developed system. In Table 1 the key meteorological variables and their respective specifications are described:

Sensor Calibration was performed in a controlled laboratory environment. The readings from each sensor were periodically compared against certified standard reference instruments to minimize measurement errors and ensure accuracy throughout the operational period.

Table 1: *Static and Dynamic Characteristics of Weather Station Sensors.*

Variable	Sensor Model	Range	Accuracy	Interface
Temperature	SHT30	–40 °C a 125 °C	±0.3 °C	I2C
Relative Humidity	SHT30	0% a 100% HR	±2% HR	I2C
Atmospheric Pressure	BME280	300 a 1100 hPa	±1 hPa	I2C
Wind Speed	SKU 6410 (Magnetic Sensor)	0 a 60 m/s	±0.5 m/s	Digital/Analog
Wind Direction	SKU 6410 (Potentiometer/wind vane)	8 cardinal directions	-	Analog
Precipitation	SKU 6466 (Tilting bowl rain gauge)	-	0.2 mm per pulse	Digital
Solar Radiation	SKU 6450 (Photovoltaic Pyranometer)	-	Irradiance in W/m²	Analog

Table 2: *Data Acquisition Phases for the Weather Station Prototype.*

Phase	Duration	Temporal Range	Functional Purpose
External Historical Baseline	~44 Months	Jan 2021 – Aug 2024	Primary training of the LSTM model using the Jena Climate Dataset (400,000 records).
Initial Supplemental Collection	3 Months	Sept 2021 – Dec 2021	Private data acquisition for model refinement and preliminary hardware verification.
Sensor Stability Validation	~4 Months	Sept 2021 – Jan 2022	High-frequency sampling (20,000 readings) to validate sensor consistency.
Operational Validation	6 Months	Jul 2024 – Dec 2024	Continuous field testing of the complete IoT-ML pipeline and prediction accuracy.
Predictive Performance Evaluation	12 Months	Aggregate Annual Cycle	Collection of 20,000+ samples used for final metric calculations.

Data Acquisition and Communication Protocols

Data Acquisition and Transmission: The data acquisition software was implemented in C++ on the Arduino™ UNO. This algorithm periodically reads and processes the sensor data, temporarily storing it in the microcontroller's internal memory. The data is then structured into a 40-byte byte array, representing 10 float values, and transmitted wirelessly via the LoRa™ module operating in TX (Transmitter) mode. The firmware utilizes standard libraries, including Adafruit_Sensor, Adafruit_BME280, and RH_RF95, to manage sensor readings and communication protocols.

Data Reception and Storage: The Raspberry® Pi 3 operates a Python™ script in RX (Receiver) mode, relying on libraries such as SX127x.LoRa, struct, pandas, datetime, and MySQLdb to receive and unpack the binary LoRa™ payload into float variables. A critical step involves initial pre-processing, including filtering and cleaning of outliers, to ensure data integrity. The processed data is stored using a dual-strategy approach. For long-term archival and retrospective analysis, the data is organized into a DataFrame using the pandas library and stored in .csv files. Simultaneously, the data is written to a MySQL relational database for real-time visualization, dynamic access, and use by the predictive models. This comprehensive system operated continuously over a six-month experimental period, providing consistent and traceable historical data.

Predictive Modeling and Analysis

Data Sources and Preprocessing: The system's predictive capacity was achieved through the implementation of a multivariable LSTM neural network. The model training utilized a large historical time series dataset derived from two sources: The publicly available Jena Climate Dataset, containing approximately 400,000 records of 14 meteorological parameters spanning from January 2021 to August 2024.

The data architecture of this study is built upon a stratified temporal framework that distinguishes among foundational historical training data, hardware-specific supplemental data, and operational validation phases. The prototype weather station performed continuous data acquisition over an initial three-month period (September–December 2021), during which high-frequency sensor measurements were logged to supplement publicly available historical datasets, yielding localized meteorological variables under field conditions.

To enable comprehensive model training and validation, the collected prototype data were integrated with historical records into a combined dataset covering approximately twelve months. Subsequent field validation and system testing extended over a six-month timeframe, allowing an in-depth evaluation of both hardware performance and predictive model accuracy across diverse climatic scenarios. This temporal framework provided the necessary microclimatic variability required to train the multivariate LSTM model and assess its real-time forecasting capabilities. Table 2 describes the time frame for the data recorded in the prototype of weather station.

During the initial three-month operational period of the prototype weather station, continuous data acquisition at a sampling frequency of one measurement every 10 seconds yielded approximately 777,600 records per sensor. Across the full

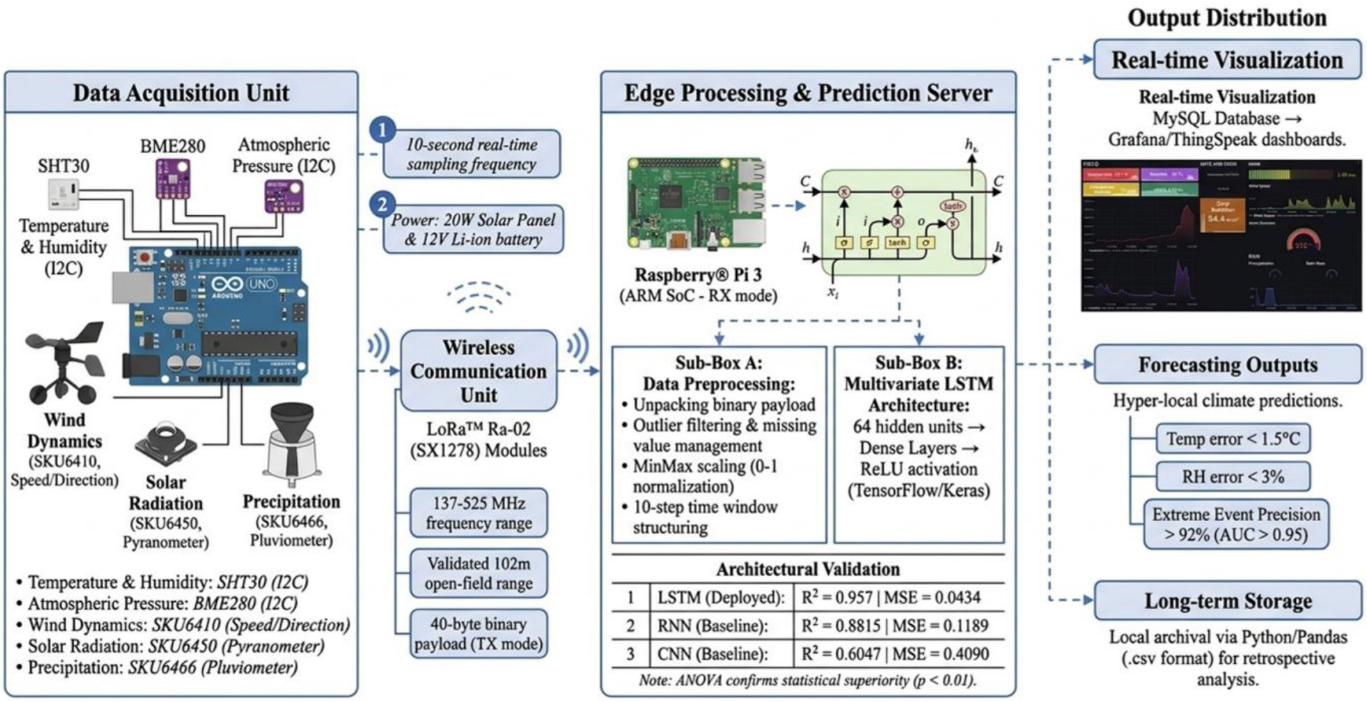

Fig. 2: Smart Weather Station Workflow, from IoT Acquisition to Deep Learning Prediction.

suite of monitored meteorological variables such as: temperature, relative humidity, atmospheric pressure, wind speed and direction, precipitation, and solar radiation. This represented over 5.4 million data points for the training phase. By effectively capturing rapid microclimatic fluctuations, this high-density dataset significantly enhanced temporal granularity when integrated with external historical records.

Supplemental private data collected directly by the prototype station over an initial three-month operational period. The combined dataset underwent a rigorous preprocessing pipeline prior to training. This involved MinMax scaling (normalization to the range 0-1), the identification and removal of outliers, and the management of missing values. The predictive framework was developed using a high temporal resolution, with data acquired at 10-second intervals to preserve microclimatic fluctuations. The complete dataset (comprising at least 400,000 records from the Jena Climate Dataset supplemented by local prototype observations) was partitioned chronologically into training (80%), validation (10%), and testing (10%) sets to maintain temporal consistency and prevent data leakage.

The LSTM model used input sequences with a lookback window of 10 time steps (100 seconds). Preliminary experiments demonstrated that this window length provided an optimal balance between capturing short-term temporal dependencies and computational efficiency. Concerning the validation methodology, traditional k-fold cross-validation was avoided because random data shuffling violates the sequential structure of time series, leading to data leakage. Instead, chronological splitting (temporal cross-validation) was applied to preserve the time-series order, ensuring a realistic evaluation of forecasting performance on unseen future data.

LSTM Architecture and Training: The LSTM network architecture was composed of a single LSTM layer featuring 64 hidden units, followed by two fully connected (dense) layers. The choice to utilize a single-layer LSTM architecture with 64 hidden units was a deliberate design decision based on a systematic empirical evaluation. Preliminary experiments compared single-layer and multi-layer LSTM configurations in terms of predictive accuracy, computational efficiency, and overfitting tendency on the validation dataset. The single-layer model consistently achieved comparable or superior performance with reduced computational complexity and lower risk of overfitting compared to deeper architectures.

The ReLU activation function was applied to introduce non-linearity. The model was implemented using the TensorFlow 2.x library with the Keras API in a Python™ environment. The training process was executed using the Adam (Adaptive Moment Estimation) optimizer, selected for its adaptive learning rate capabilities essential for handling the non-stationary dynamics of microclimate data. The learning rate ($\eta = 10^{-4}$) was determined through empirical hyperparameter tuning, ensuring monotonic loss decay while preventing premature convergence to suboptimal local minima when processing the 10-step sequence windows. The Mean Squared Error (MSE) was used as the loss function, and early stopping was implemented to prevent overfitting.

To significantly reduce computational time during model development, the initial training was performed on a high-performance machine (Intel Core i5-11300H) before the finalized weights were deployed onto the Raspberry® Pi 3 for real-time inference. Final predicted values were subjected to a denormalization step to revert them to their original physical units for meaningful interpretation.

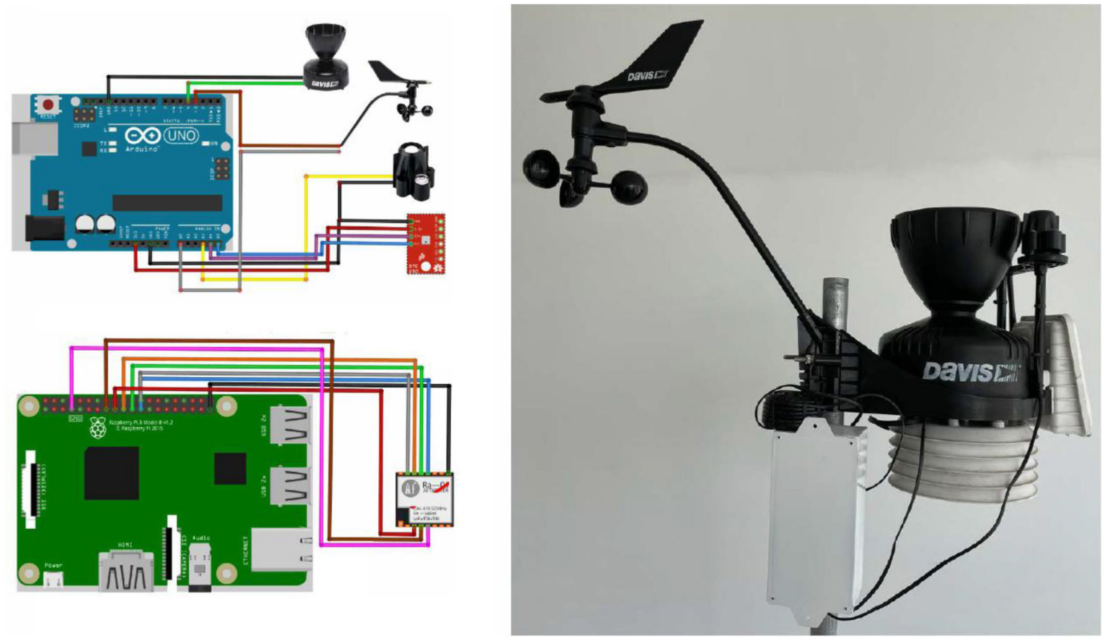

Fig. 3: Weather Station Designed and Implemented

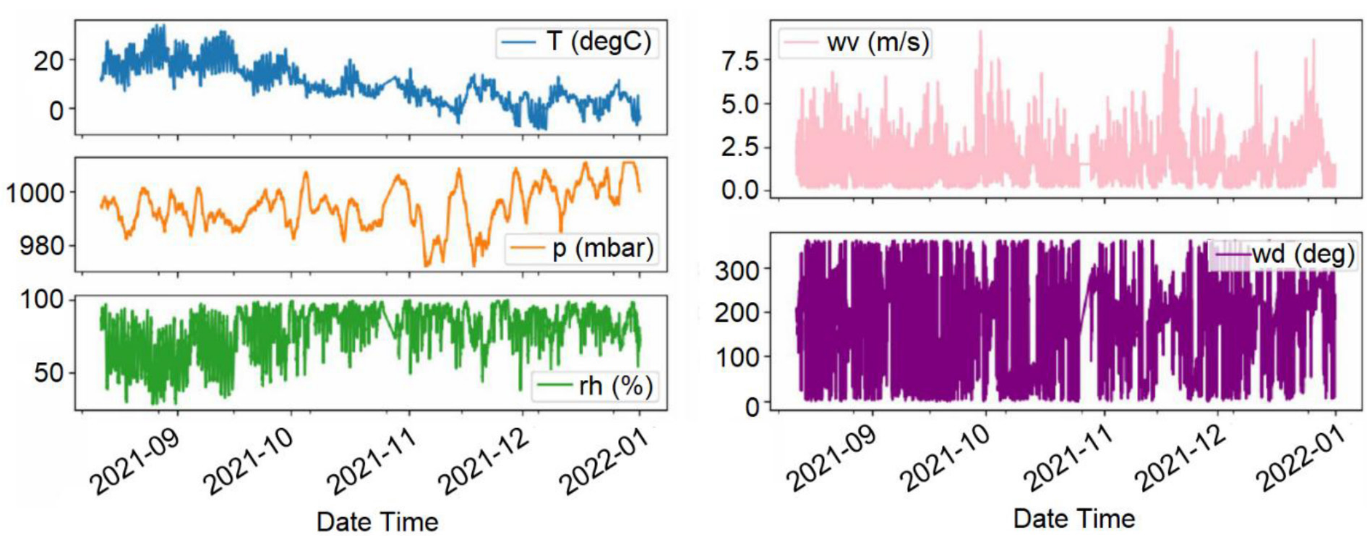

Fig. 4: Record of meteorological variables of T, p, rh, wv and wd between 2021-09 and 2021-12.

System Validation and Quality Assurance

The complete system, including both hardware stability and predictive performance, was evaluated through a six-month protocol of real-world testing. Throughout this period, the predictions generated by the LSTM model were continuously compared against the real-time sensor measurements from the deployed station. Quality assurance procedures included the implementation of anomaly detection algorithms to automatically reject sensor readings outside of physically plausible ranges. Additionally, robustness tests were executed to assess the station's stability and the data coherence under adverse conditions, such as sudden wind shifts and heavy precipitation events.

Performance metrics were calculated on the independent test set, including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2). The model demonstrated satisfactory predictive precision, achieving a mean error of less than 1.5 °C for temperature and 3% RH for relative humidity during cross-validation. The methodology, supported by detailed technical specifications, PCB designs, and comprehensive software documentation (including firmware and Python™ scripts), ensures the technical reproducibility of the entire system for future implementation in alternative contexts. Fig. 2 presents the workflow schematic, illustrating the overall process from data acquisition and transmission to model development and prediction.

RESULTS AND DISCUSSION

System Implementation and Data Acquisition Validation

The integrated, autonomous meteorological station was

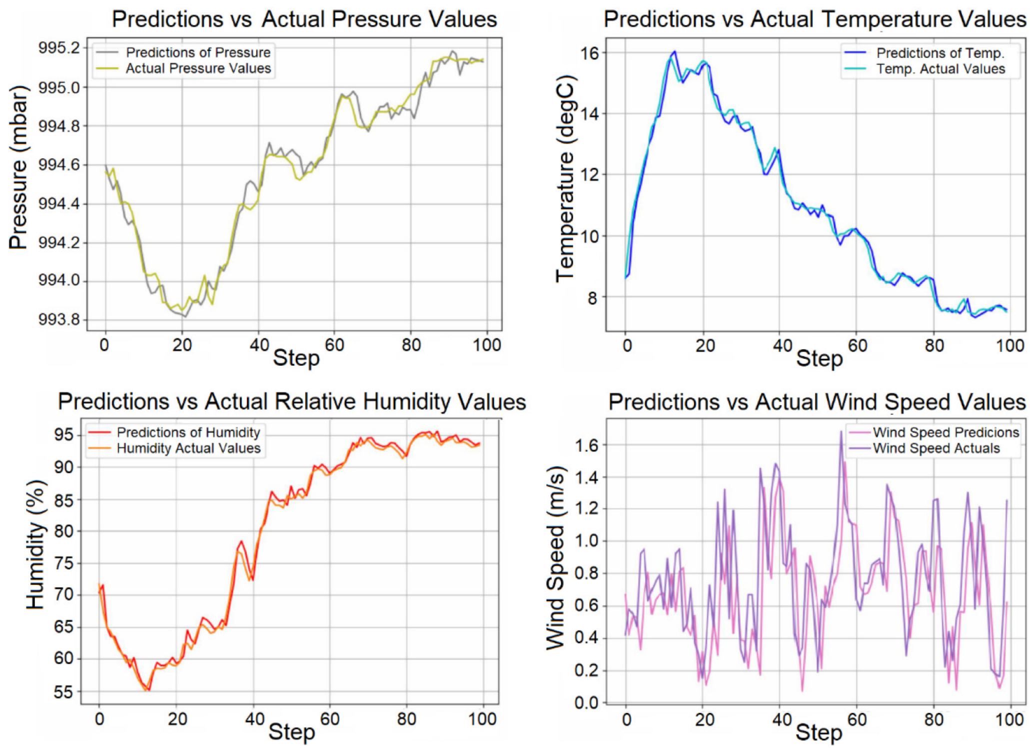

Fig. 5: Record between predicted and actual meteorological variables with LSTM model.

Table 3: Quantitative Results Summary (RNN vs. LSTM vs. CNN).

Metric / Component	Value Obtained	Context and Observation
LSTM Regression Performance		
Coefficient of Determination (R^2)	0.957	Indicates an exceptional fit, very close to perfection for meteorological time series.
Mean Squared Error (MSE)	0.0434	Significantly exceeds the project's design objective (MSE < 0.05).
Mean Temperature Error (T)	< 1.5 °C	Error achieved during cross-validation.
Mean Relative Humidity Error (RH)	< 3% RH	Error achieved during cross-validation.
Model Comparison (Test Set)		*LSTM surpassed comparison models with statistical significance ($p < 0.01$).*
RNN R^2/ MSE	0.8815 / 0.1189	Performance of the traditional Recurrent Neural Network model.
CNN R^2/ MSE	0.6047 / 0.4090	Performance of the Convolutional Neural Network model.
Model Robustness and Generalization		Evaluation of classification capability for extreme events.
Precision	> 92%	Ability to correctly identify a true positive extreme event.
Recall	89%	Ability to capture most actual extreme events.
Area Under the Curve (AUC)	> 0.95	Confirms high discriminative power of the model.
Cross-Validation	10-Fold	Minimal variability in metrics, confirming low propensity for overfitting.

successfully designed and deployed following the architecture detailed in the Methods section (Fig. 2). The physical implementation was verified for structural integrity, power autonomy, and sensor housing efficiency against environmental factors. Fig. 3 illustrates the design and implementation of: the sensor data acquisition unit using the Arduino™ UNO board, the processing unit with the Raspberry® Pi, the remote communication unit with the LoRa™ Module, and the weather station to be installed in the field.

The system's dual storage mechanism, utilizing local *.csv archival and real-time synchronization to a MySQL relational database, ensured data consistency and traceability throughout the six-month experimental period. During the initial operational phase, time series data were successfully conditioned and captured for all monitored meteorological variables. A sample period encompassing over 20,000 readings validated the consistency and temporal evolution of temperature (T (degC) (°C)), atmospheric pressure (p (mbar)), relative humidity (rh (%)), wind speed (wv (m/s)), and wind direction (wd (deg)) (See Fig. 4).

Data records were confirmed to be stable and consistent across all variables prior to input into the predictive modeling pipeline.

The hardware validation focused on verifying the precision of the integrated sensors and the reliability of the wireless communication link. Comparative measurements between the deployed prototype and two external reference instruments (a commercial Davis Vantage Pro 2 station and a manual UPTC station), demonstrated a minimal error margin across all measured

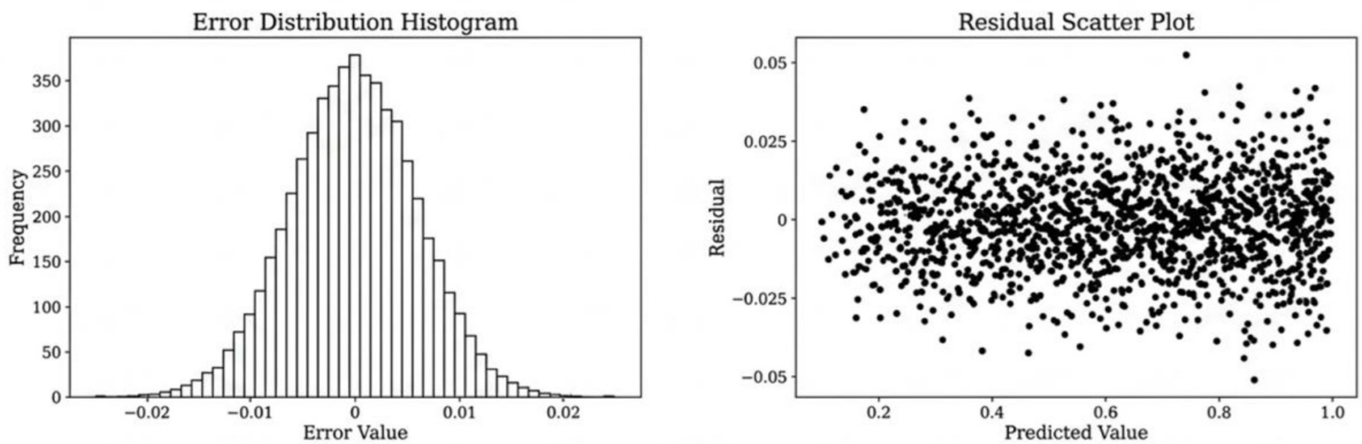

Fig. 6: Robustness Evaluation and Residual Analysis.

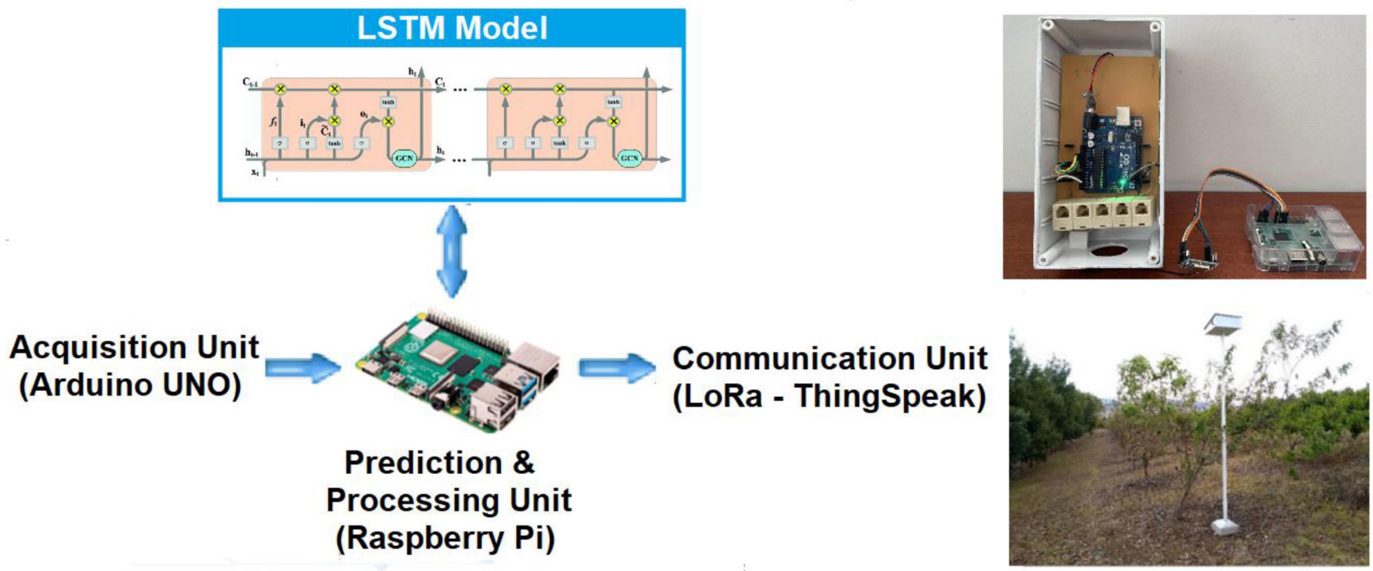

Fig. 7: Implementation and installation of the smart weather station in the field.

variables. The LoRa™ Ra-02 communication link was functionally validated, sustaining reliable data transmission up to a range of 102 meters in an open field environment. This confirmed the system's viability for deployment in remote agricultural and environmental monitoring applications where traditional infrastructure is limited.

Comparative Analysis of Predictive Model Performance

The predictive core of the system, a multivariable LSTM neural network, was rigorously evaluated using a test set derived from a dataset of over 20,000 samples collected over a 12-month period. The model achieved high predictive accuracy, evidenced by a R^2 of 0.957 and a MSE of 0.0434. This performance significantly surpassed the project's target of maintaining the MSE below 0.05, suggesting exceptional model fit.

To contextualize the architectural superiority of the LSTM approach, its performance was benchmarked against alternative sequential models: a traditional Recurrent Neural Network (RNN) and a Convolutional Neural Network (CNN). The comparative results are summarized in Table 3: the RNN model yielded an R^2 of 0.8815 and an MSE of 0.1189, while the CNN demonstrated the lowest performance, with an R^2 of 0.6047 and an MSE of 0.4090. A one-way Analysis of Variance (ANOVA) conducted on the predictive residuals confirmed the statistical difference in performance, with the LSTM model demonstrating a highly significant advantage over both comparator models ($p < 0.01$).

The superior performance of the LSTM architecture is attributed to its specialized recurrent structure, which effectively addresses the long-term temporal dependencies and non-stationarity characteristic of complex meteorological time series data. This capability was visually confirmed by the near-perfect alignment observed between the LSTM-predicted values and the actual measurements for stable variables like temperature, pressure, and humidity (Fig. 5). While wind speed and direction inherently exhibit higher volatility, the LSTM successfully captured the overall trend and magnitude of their variations.

Model Robustness and Generalization Assessment

The model's generalization capability was confirmed

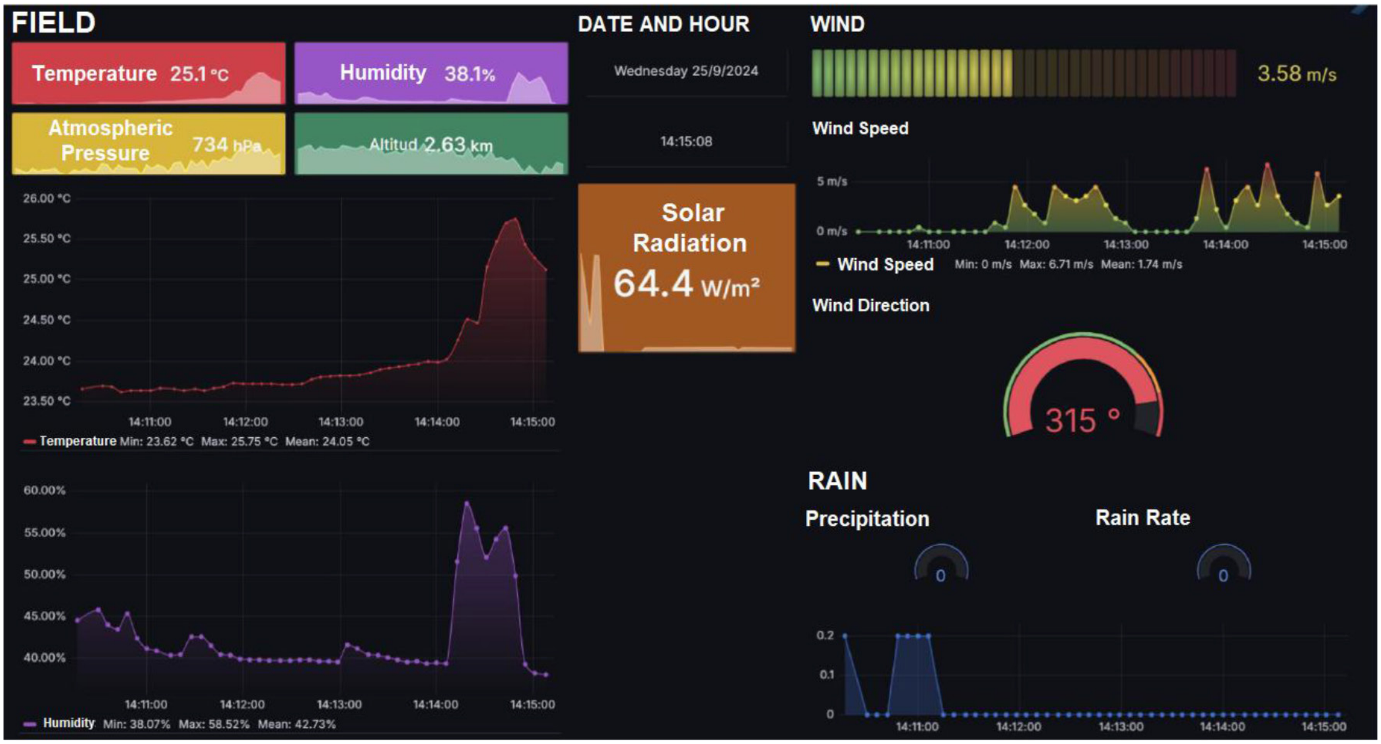

Fig. 8: Visualization and Analysis of Meteorological Variables in Remote Station.

through 10-fold cross-validation, which yielded minimal variability in performance metrics across all folds, suggesting high stability and low risk of overfitting. Furthermore, the model's robustness was assessed by evaluating its discriminative power for classification tasks related to extreme weather event forecasting (e.g., intense rain or high solar radiation).

The LSTM model achieved a Precision exceeding 92% and a Recall of 89% in identifying significant environmental events, demonstrating strong discriminative power. This reliability in classifying extreme events was further validated by Receiver Operating Characteristic (ROC) curves, which consistently resulted in an Area Under the Curve (AUC) greater than 0.95. The statistical reporting was enhanced by the inclusion of 95% confidence intervals (CI), which confirmed the consistency and statistical reliability of the reported performance metrics across the validation subset. Rigorous data quality assurance protocols, including outlier detection via standard deviation analysis and residual analysis, were instrumental in ensuring that the reported findings were objective and free from sampling or measurement biases.

The robustness of the LSTM model was evaluated through error distribution profiling and residual analysis (Fig. 6). The error histogram exhibits a normal distribution centered near zero, while the residual scatter plot confirms homoscedasticity across the range of predicted values. These statistical properties demonstrate that the model effectively captures non-linear temporal dependencies without systematic bias, thereby ensuring reliable forecasts for microclimate decision-making.

Discussion: Operability, Contextualization, and Novelty

The integration of the trained LSTM algorithm onto the Raspberry® Pi 3 for local inference establishes the system as an efficient and scalable solution for edge computing in environmental science. By offloading the computationally demanding training phase to a more powerful external machine (Intel Core i5-11300H), the design optimizes the trade-off between model precision and system portability, enabling the small-form-factor Raspberry® Pi to execute complex real-time predictions efficiently. This is a significant advance over studies relying solely on cloud-based inference, offering enhanced reliability and reduced latency crucial for timely decision-making.

The field deployment demonstrated high operability, with the system's architecture allowing for seamless data flow from acquisition to visualization (See Fig. 7).

The integration with a MySQL database provides a structured backend for data organization, while the use of Grafana enables flexible, real-time visualization and alerting capabilities (Fig. 8). This complete technological pipeline addresses a key gap in existing literature by offering a portable solution that cohesively integrates local acquisition, deep learning-based prediction, and sophisticated visualization within a single autonomous unit.

Visual assessment of the real-time prediction logs (Figs. 9, 10, 11, 12, 13, and 14), which recorded predictions every 10 minutes over four months, between each data point from August to December 2024, confirmed the ability of the synthesized LSTM model to accurately estimate general trends and variations within average meteorological ranges. The consistency between the graphical logs and the quantitative metrics (MSE, R^2) further validated the model's practical utility for real-life climate decision-making.

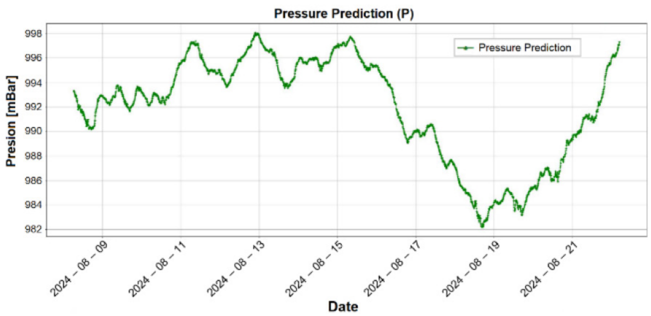

Fig. 9: Record of predictions of Atmospheric Pressure as a function of time.

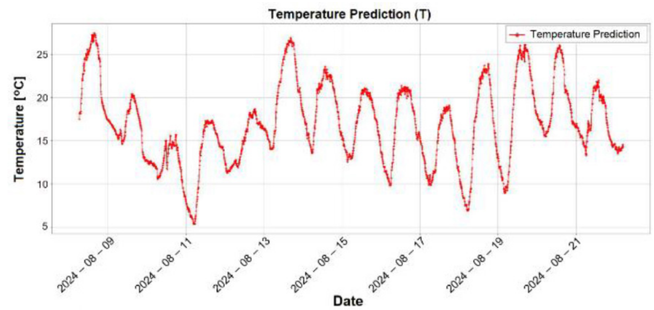

Fig. 10: Record of predictions of Ambient Temperature as a function of time.

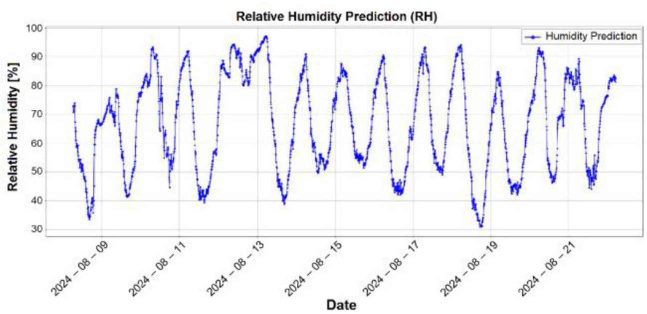

Fig. 11: Record of predictions of Relative Humidity as a function of time.

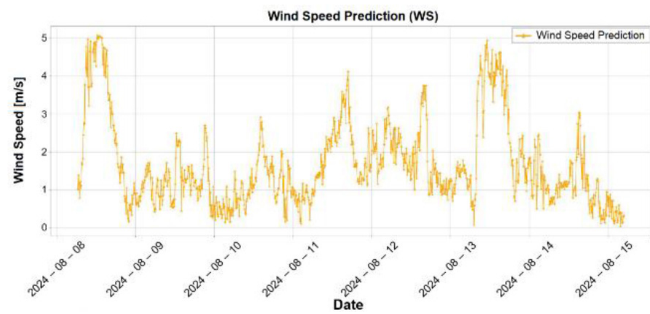

Fig. 12: Record of predictions of Wind Speed as a function of time.

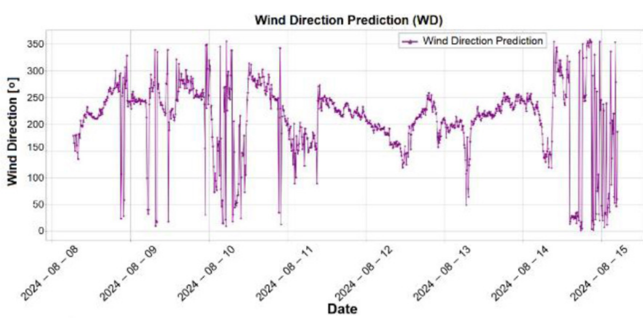

Fig. 13: Record of predictions of Wind Direction as a function of time.

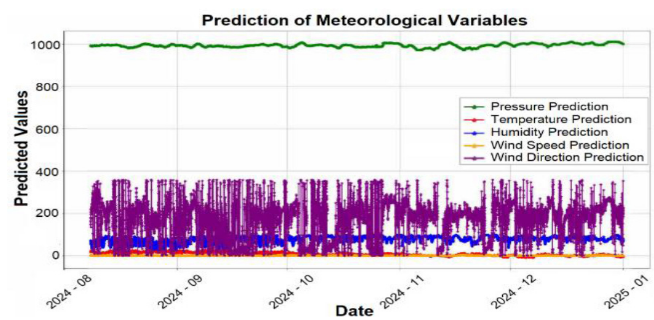

Fig. 14: Record of predictions of meteorological variables as a function of time.

The primary novelty of this work lies in the demonstrably superior predictive performance of the optimized LSTM model synthesized for a low-power edge device, coupled with the proven long-range capability of the LoRa™ communication protocol. While the system's reliance on pre-trained weights is an optimization for portability, a limitation exists in the computational resources of the Raspberry Pi® 3, which restricts the capacity for continuous, on-device model retraining. Future research will explore the potential for implementing lightweight, incremental learning algorithms directly on the edge device to further enhance the system's long-term adaptability to evolving microclimatic conditions.

Validating low-cost, decentralized "Smart Weather Station" architectures demonstrates that open-hardware components (e.g., ESP8266/NodeMCU with BME280 sensors) can achieve scientific-grade accuracy. Recent benchmarks confirm high fidelity across microclimate variables, notably Azevedo et al., (2025) achieving $R^2 = 0.93$ (MAE = 0.30 mm, c = 0.9202) for evapotranspiration estimation and Yang et al., (2024) obtaining $R^2 = 0.996$ for temperature modeling. However, component selection is critical, as high BME280 reliability (RMSE < 0.25 °C) contrasts with low-performing sensors (e.g., H202L, TSL2561).

Crucially, these systems offer viable alternatives to expensive commercial stations (e.g., Campbell Scientific units costing $4,500 USD) democratizing high-precision environmental sensing. By integrating IoT protocols (MQTT, Adafruit IO), local SPIFFS storage, and NTP synchronization, these architectures bridge the "Airport-to-Local" microclimate discrepancy. This localized precision is essential for microscale modeling; for instance, site-specific data shifted building energy estimates by a +2% to +14% increase in cooling and a –1% to –10% reduction in heating requirements. Ultimately, combining edge intelligence with low-cost hardware establishes a scalable benchmark for environmental informatics, enhancing resource efficiency in precision agriculture and urban climate monitoring.

CONCLUSION

This research successfully designed and implemented an autonomous, IoT-enabled smart meteorological station that integrates high-frequency data acquisition with Deep Learning-based prediction capabilities. The results validate the effectiveness of the system at both the hardware and predictive modeling levels.

The prototype demonstrated data acquisition precision and reliability comparable to high-end commercial stations. This was confirmed by a minimal error margin when comparing its measurements against a reference station, the Davis Vantage Pro 2, and a manual UPTC station. Operational validation verified the system's integrity in the field over a six-month period and confirmed the LoRa™ functionality for reliable, long-range data transmission (up to 102 meters) in environments lacking traditional network infrastructure.

Regarding predictive capability, the LSTM model consistently outperformed comparative architectures, including RNNs and CNNs. The LSTM model achieved a R^2 of 0.957 and a MSE of 0.0434, confirming a robust and precise capability to capture and forecast the inherent non-linear temporal dependencies within meteorological time series.

Additionally, the model exhibited high generalization and robustness, achieving a Precision exceeding 92% for the classification of extreme climatic events. The primary scientific contribution of this work lies in the cohesive integration of a high-precision Deep Learning model onto a low-power edge device (Raspberry® Pi 3). This architecture resolves a critical limitation identified in the literature—the scarcity of portable and accessible solutions that combine real-time data collection with hyper-local AI-based forecasting. By synthesizing the LSTM algorithm for local inference, the system optimizes the trade-off between computational precision and portability, providing an efficient, scalable, and low-latency alternative for climatic monitoring in precision agriculture, risk management, and environmental studies in remote areas.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the technological support provided by the I^{2}E Research Group, Electronic Engineering Program, at the Universidad Pedagógica y Tecnológica de Colombia (Tunja, Colombia), for the development of the study presented within the framework of Project SGI 3760: "Desarrollo de una Estación Meteorológica orientada a la Predicción de Variables Climáticas mediante la Aplicación de Técnicas de Inteligencia Artificial."

Furthermore, the authors thank the Macrypt Research Group at the Universidad de Los Llanos (Villavicencio, Colombia), for their technological and academic assistance during the execution of the research project: "Performance monitoring system for citrus crops using analysis of information acquired by remote sensors" (Code: C09-F01-020-2022 FCBI).

Author's certificate: The manuscript or its part is not under consideration for publication elsewhere and the same has been approved by all co-authors.

Conflict of Interests: The authors declare that there is no conflict of interest regarding the publication of this article.

Funding: No funding was received for this work.

Data Availability: Data can be provided on request.

Author's contribution: **Jenny-A. Rosales-Agredo**: Conceptualization, Writing original draft, Data collection, Validation, Methodology, Investigation, Formal analysis; **Fabián-R. Jiménez-López:** Conceptualization, Writing original draft, Data collection, Validation, Methodology, Investigation, Formal analysis; **Andrés-F. Jiménez-López:** Conceptualization, Reviewing original draft, Supervision.

Disclaimer: The contents, opinions, and views expressed in the research article published in the Journal of Agrometeorology are the views of the authors and do not necessarily reflect the views of the organisations they belong to.

Publisher's Note: The periodical remains neutral concerning jurisdictional claims in published maps and institutional affiliations.

REFERENCES

Alam, M., Islam, M., Nayan, N., & Uddin, J. (2025). An IoT based real-time environmental monitoring system for developing areas. *Journal of Advanced Research in Applied Sciences and Engineering Technology, 52*(1), 106–121. https://doi.org/10.37934/araset.52.1.106121

Alves, D., Mendonça, F., Mostafa, S. S., & Morgado, F. (2024). A comprehensive IoT cloud-based wind station ready for real-time measurements and artificial intelligence integration. *E-Prime: Advances in Electrical Engineering, Electronics and Energy, 10*, 100862. https://doi.org/10.1016/j.prime.2024.100862

Asanza, V., Pico, R. E., Torres, D., Santillan, S., & Cadena, J. (2021). FPGA-based meteorological monitoring station. In *2021 IEEE Sensors Applications Symposium (SAS)* (pp. 1–5). IEEE. https://doi.org/10.1109/SAS51076.2021.9530151

Azevedo, A. T., Coelho, R. D., Carnevskis, E. L., de Almeida, A. M., & Tabile, R. A. (2025). Development of an automatic weather station for irrigation management via IoT. *Agricultural Research, 14*(3), 529–538. https://doi.org/10.1007/s40003-024-00786-8

Bhatia, S., Naib, B. B., Goel, N., Agarwal, B., Mallikarjun, B., & Arvindhan, M. (2024). A new Era in weather forecasting: Harnessing the potential of machine learning. In *2024 International Conference on Computing, Power, and Communication Technologies (IC2PCT)* (pp. 1355–1361). IEEE. https://doi.org/10.1109/IC2PCT60090.2024.10486245

Bhattacharjee, S., Biwalkar, N., Sur, K., Kaur, S., & Singh, S. P. (2025). Application of artificial intelligence and statistical recurrent models in predicting rainfall: A case study of Ludhiana, Punjab. *Journal of Agrometeorology, 27*(3), 395–398. https://doi.org/10.54386/jam.v27i3.3038

Budjač, R., & Gašpar, G. (2024). Proposal for a weather monitoring system utilizing IoT technologies. In *Proceedings of the 35th Central European Conference on Information and Intelligent Systems (CECIIS)* (pp. 67–74).

Eleftheriou, A., Karvelis, P., Kouvaris, K., & Stylios, C. (2018). Microclimate prediction utilising machine learning approaches. In *International Workshop on Metrology for the Sea (MetroSea)* (pp. 198–200). IEEE. https://doi.org/10.1109/MetroSea.2018.8657903

Faid, A., Sadik, M., & Sabir, E. (2022). An agile AI and IoT-augmented smart farming: A cost-effective cognitive weather station. *Agriculture, 12*(1), Article 35. https://doi.org/10.3390/agriculture12010035

Gao, P. (2025). Intelligent disaster prevention and mitigation system based on machine learning algorithm. *Procedia Computer Science, 262*, 1374–1380. https://doi.org/10.1016/j.procs.2025.05.184

Grekhnev, I., Shchelkanov, M., & Matrenin, P. (2023). Comparative analysis of machine learning algorithms using weather station grid short-term load forecasting. In *Belarusian-Ural-Siberian Smart Energy Conference (BUSSEC)* (pp. 60–64). IEEE. https://doi.org/10.1109/BUSSEC59406.2023.10296417

Hachimi, C., Belaqziz, S., Khabba, S., Sebbar, B., Dhiba, D., & Chehbouni, A. (2023). Smart weather data management based on artificial intelligence and big data analytics for precision agriculture. *Agriculture, 13*(1), Article 95. https://doi.org/10.3390/agriculture13010095

Herdianzenda, L., Siswanto, D., & Istiadi, I. (2021). Design of portable automatic weather station based on Raspberry Pi. *Journal of Science and Applied Engineering, 4*(1), 21–31. https://doi.org/10.31328/jsae.v4i1.3048

Hilda, J. J., Jinny, S. V., Srilekha, P. M., & Selvakasi, V. (2025). IoT-based smart weather monitoring system. In *2025 6th International Conference on Innovative Trends in Information Technology (ICITIIT)* (pp. 1–6). IEEE. https://doi.org/10.1109/ICITIIT64777.2025.11041600

Huang, Z. Q., Chen, Y. C., & Wen, C. Y. (2020). Real-time weather monitoring and prediction using city buses and machine learning. *Sensors, 20*(18), Article 5173. https://doi.org/10.3390/s20185173

Hussein, Z. K., Hadi, H. J., Mutaleb, M. R., & Mezaal, Y. S. (2020). Low-cost smart weather station using Arduino and ZigBee. *TELKOMNIKA (Telecommunication Computing Electronics and Control), 18*(1), 282–288. https://doi.org/10.12928/TELKOMNIKA.v18i1.12784

Ioannou, K., Karampatzakis, D., Amanatidis, P., Aggelopoulos, V., & Karmiris, I. (2021). Low-cost automatic weather stations in the Internet of Things. *Information, 12*(4), Article 146. https://doi.org/10.3390/info12040146

Islam, K., Chowdhury, F., Farhan, K., Uddin, M. R., & Hasan, M. (2024). IoT integrated air quality and weather monitoring system using Android application. In *International Conference on Recent Progresses in Science, Engineering and Technology (ICRPSET)* (pp. 1–8). IEEE. https://doi.org/10.1109/ICRPSET64863.2024.10955912

Jolly, F., Uddin, G., Alim, M., Kumar, R., Dutta, A., Reya, M., & Tasnim, N. (2024). Analyzing the efficiency of Arduino UNO microcontroller in monitoring and controlling the microclimatic parameters of greenhouse. *Journal of Agrometeorology, 26*(1), 51–55. https://doi.org/10.54386/jam.v26i1.2520

Kalimuthu, M., Vaishnavi, P., & Kishore, M. (2020). Crop prediction using machine learning. In *2020 3rd International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 926–932). IEEE. https://doi.org/10.1109/ICSSIT48917.2020.9214190

Luo, Y., Zhang, Y., Huang, Y., & Tong, S. (2023). Low-power LoRa weather station with adaptive data fusion for agriculture. In *2023 8th International Conference on Intelligent Computing and Signal Processing (ICSP)* (pp. 1077–1081). IEEE. https://doi.org/10.1109/ICSP58490.2023.10248432

Math, R. K., & Dharwadkar, N. V. (2018). IoT-based low-cost weather station and monitoring system for precision agriculture in India. In *International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)* (pp. 81–86). IEEE. https://doi.org/10.1109/I-SMAC.2018.8653749

Megantoro, P., Aldhama, S. A., Prihandana, G. S., & Vigneshwaran, P. (2021). IoT-based weather station with air quality measurement using ESP32 for environmental aerial condition study. *TELKOMNIKA (Telecommunication Computing Electronics and Control), 19*(4), 1316–1325. https://doi.org/10.12928/TELKOMNIKA.v19i4.18990

Nasrat, M. A. (2025). Implementation of a weather station and forecasting weather using IoT: A review. In *2025 28th International Conference on Soft Computing and Measurements (SCM)* (pp. 290–295). IEEE. https://doi.org/10.1109/SCM66446.2025.11060292

Ngebani, I., Nfanyana, K., Sheikh, S., & Ishmael, Z. (2018). Design of a microcontroller-based Internet of Things weather station. *Journal of Engineering Research and Application, 8*(8), 28–33. https://doi.org/10.9790/9622-0808012833

Omar, N., Al-Zebari, A., & Sengur, A. (2021). Deep learning approach to predict forest fires using meteorological measurements. In *2021 2nd International Informatics and Software Engineering Conference (IISEC)* (pp. 1–4). IEEE. https://doi.org/10.1109/IISEC54230.2021.9672446

Perera, K. P., Lakmal, H. K., Fernando, K. J., & Nirmal, W. C. (2025). Hybrid deep learning for portable weather forecasting: Real-time predictions using CNN-LSTM-transformer models. In *International Research Conference on Smart Computing and Systems Engineering (SCSE)* (pp. 1–6). IEEE. https://doi.org/10.1109/SCSE65633.2025.11031062

Plein, M., Feigel, G., Zeeman, M., Dormann, C., & Christen, A. (2025). Using gradient boosting for gap-filling to analyze temperature and humidity patterns in an urban weather station network in Freiburg, Germany. *Urban Climate, 62*, Article 102496. https://doi.org/10.1016/j.uclim.2025.102496

Roy, D. S. (2020). Forecasting the air temperature at a weather station using deep neural networks. *Procedia Computer Science, 178*, 38–46. https://doi.org/10.1016/j.procs.2020.11.005

Sakthipriya, D., & Chandrakumar, T. (2024). Weather-based paddy yield prediction using machine learning regression algorithms. *Journal of Agrometeorology, 26*(3), 344–348. https://doi.org/10.54386/jam.v26i3.2598

Wicaksana, H., Ananda, N., Budiawan, I., Santoso, B., Handoko, R., Maulana, A., Suciarti, S., & Utoro, A. (2022). Air temperature sensor estimation on automatic weather station using ARIMA and MLP. *Ultimatics: Jurnal Teknik Informatika, 14*(2), 103–110. https://doi.org/10.31937/ti.v14i2.2865

Yang, S., Zhan, D., Stathopoulos, T., Zou, J., Shu, C., & Wang, L. L. (2024). Urban microclimate prediction based on weather station data and artificial neural network. *Energy and Buildings, 314*, 114283. https://doi.org/10.1016/j.enbuild.2024.114283

Zhang, H., Liu, Y., Zhang, C., & Li, N. (2025). Machine learning methods for weather forecasting: A survey. *Atmosphere, 16*(1), Article 82. https://doi.org/10.3390/atmos16010082