SOIL QUALITY MONITORING AND EVALUATION SYSTEM USING MACHINE LEARNING AND LORAWAN WIRELESS COMMUNICATION

Van Lic Tran, Tran Le Xuan Huy, Nguyen Dai, Bao Ngan Truong Nguyen, Le Hong Nam*

The University of Danang - University of Science and Technology, Vietnam

*Corresponding author: lhnam@dut.udn.vn

(Received: April 05, 2025; Revised: June 05, 2025; Accepted: June 15, 2025)

DOI: 10.31130/ud-jst.2025.23(9C).545E

Abstract - This study demonstrates the successful deployment and operation of LoRaWAN technology, yielding significant positive outcomes. The system effectively collects critical soil parameters, including moisture, temperature, electrical conductivity (EC), pH, and NPK levels, thereby fulfilling the requirements for agricultural soil monitoring and management. Leveraging the LoRaWAN protocol, the system ensures reliable, long-range data transmission with minimal energy consumption, making it highly suitable for remote farmlands where traditional connectivity is limited. Moreover, its modular architecture and scalability provide flexibility for large-scale deployment across diverse agricultural regions. In addition, the collected data can be integrated with machine learning models to analyze soil dynamics, predict future trends, and optimize resource allocation, such as fertilizer and irrigation scheduling. Ultimately, the system contributes to sustainable farming practices by recommending suitable crops and improving overall agricultural productivity.

Key words - Soil Quality; LoRaWAN; Machine Learning; NPK; PH; electrical conductivity (EC).

1. Introduction

Soil quality is vital for optimizing crop yield and quality, as it supplies plants with essential nutrients and water. However, assessing and monitoring soil quality presents significant challenges. Inadequate infrastructure for soil monitoring, particularly in underdeveloped regions, limits effective assessment. Moreover, the high costs of advanced monitoring technologies often exceed the budgets of farmers and small agricultural businesses. Traditional methods, such as soil sampling and laboratory analysis, are labor-intensive, time-consuming, and unable to provide real-time or continuous data. These limitations hinder efforts to manage and enhance soil quality, ultimately reducing agricultural productivity.

Recent studies and innovations offer promising approaches to soil quality assessment. For instance, an IoT-based system for water and soil quality monitoring [1] employs sensors to measure parameters like moisture, electrical conductivity, pH, and turbidity. These sensors, embedded in soil and water, collect data that is processed by microcontrollers and transmitted via Wi-Fi to a central server for analysis. Another study [2] highlights the role of microorganisms as indicators of soil quality, given their importance in organic matter decomposition, nutrient cycling, and soil structure maintenance. In contaminated soils, bacterial community structures shift, with notable declines in biomass and activity, though diversity may remain stable. Agricultural soils also vary significantly based on soil type and land use practices.

Similarly, an IoT-based smart soil monitoring system [3] uses sensors to measure pH, temperature, and humidity, enabling farmers to select suitable crops based on soil conditions. Data is transmitted via Wi-Fi to farm managers, who receive crop recommendations through a mobile application. Another study [4] leverages remote sensing and GIS models (e.g., GEMS, RothC, CENTURY) to analyze spectral data from soil, identifying gaps in soil quality assessment and proposing improvements for agricultural fields. Additionally, a sensor-based approach using RFID technology [5] features a 13.56 MHz passive RFID tag and temperature sensors to monitor soil conditions, with data transmitted to a collection system via an RFID interrogator.

Wireless sensor networks (WSNs) also show potential for soil quality management [6]. These networks use spatially distributed sensors to monitor environmental conditions, with data exchanged between nodes until it reaches a gateway. Built on open-source platforms like Arduino and Raspberry Pi, WSNs can effectively track soil parameters. A related study [7] demonstrates the use of WSNs to monitor soil indices such as moisture, temperature, and pH.

Commercial solutions have also emerged. The Enfarm nutrient measurement kit [8] integrates sensors for NPK, temperature, and humidity, with data pre-processed for accuracy and transmitted wirelessly via Wi-Fi or Bluetooth. The ZD-2804 soil NPK meter [9] employs spectrophotometry to measure nitrogen, phosphorus, and potassium levels, offering high accuracy and ease of use. Similarly, the HI3896 NPK and pH soil test kit from Hanna Instruments [10] uses reagents to assess pH, N, P, and K levels. Despite their precision, these solutions face challenges. Many farmers lack the expertise to operate these tools or interpret results effectively. Advanced technologies are often inaccessible in rural or developing regions, and most methods focus on specific soil parameters, such as nutrients, neglecting factors like pH, moisture, soil structure, or salinity. Additionally, high costs for equipment, consumables, and maintenance pose barriers, particularly in remote areas with limited access to Wi-Fi, Bluetooth, or 4G networks, which undermines the effectiveness of data transmission systems.

To overcome these challenges, our research team proposes a soil monitoring and evaluation system integrated with LoRaWAN technology. This energy-efficient system will use long-lasting batteries and optional solar panels to reduce recharging needs. Equipped with affordable sensors to measure multiple soil parameters, the system will store and visualize data using Grafana for continuous monitoring. Furthermore, it will provide crop recommendations by comparing measured soil data with existing datasets, helping farmers select the most suitable crops for their land.

2. Soil Properties and Their Impact on Crop Growth

Soil quality is shaped by a range of environmental and nutritional factors, including temperature, moisture, electrical conductivity (EC), pH, nutrient levels (NPK), salinity, and total dissolved solids (TDS), all of which significantly influence plant growth and development. Standard soil quality requirements vary depending on the type of plant, as shown in Table 1. Temperature affects critical processes such as biochemical reactions, respiration, and photosynthesis, with each plant species requiring an optimal range. Extreme temperatures-either too high or too low-impair nutrient and water uptake, potentially leading to plant stress or death. Soil moisture determines a plant’s access to water and nutrients; insufficient moisture causes drought stress and wilting, while excessive moisture leads to waterlogging, reducing oxygen availability to roots and hindering growth.

Table 1. *Soil properties*

Temperature	Leafy vegetables	15 - 25°C
	Fruits and vegetables	20 - 25°C
Humidity	Leafy vegetables	60 - 75%
	Fruits and vegetables	60 - 70%
PH	Leafy vegetables	6.0 – 7.0 pH
	Fruits and vegetables	5.8 – 6.5 pH
EC	Leafy vegetables	1.6 – 1.8 ms/cm
	Fruits and vegetables	2.0 – 2.2 ms/cm
NPK	Leafy vegetables	N:10 - 20% P: 5 - 10% K: 10 - 20%
	Fruits and vegetables	N: 10 - 15% P: 10 - 20% K: 15 - 25%
Salt	Leafy vegetables	0.8 - 2.3 dS/m
	Fruits and vegetables	1.0 - 5.0 dS/m
TDS	Leafy vegetables	500 - 1600 mg/L
	Fruits and vegetables	700 - 3500 mg/L

Electrical conductivity (EC) serves as a measure of soil salinity, where high levels can disrupt water and nutrient absorption, causing toxicity and symptoms like wilting, yellowing leaves, and stunted growth. Soil pH governs nutrient availability, with most plants thriving within a specific pH range. Highly acidic or alkaline soils limit nutrient uptake, increasing the risk of root and leaf diseases. Macronutrients-nitrogen (N), phosphorus (P), and potassium (K)-are essential for plant health. Deficiencies or excesses of these nutrients can result in stunted growth, leaf yellowing, reduced yields, and poor fruit quality.

Soil salinity, often exacerbated by high salt content in soil or irrigation water, induces stress that lowers crop yields and disrupts physiological processes. Total dissolved solids (TDS), an indicator of salinity, influence plant growth based on the composition of dissolved ions. Effective monitoring and management of these parameters are critical for maintaining soil quality, supporting sustainable agriculture, and ensuring productive crop growth.

Sensor calibration was conducted using standard buffer solutions for pH and known reference values for NPK. Sensor measurement errors were within ±2% for EC and ±0.1 for pH readings. Repeated measurements across locations indicated low variance, confirming stability.

3. System design

Upon powering on the device, the RF Thing UCA board, TTL to RS485 module, and ES-SOIL-7IN1 sensor are activated. The UCA board sends a signal to the ground sensor, which communicates via the RS485 Modbus RTU protocol to collect data on seven soil parameters: moisture, temperature, electrical conductivity (EC), nitrogen (N), phosphorus (P), potassium (K), total dissolved solids (TDS), and salinity. The microcontroller processes the raw sensor data, performing format conversion, accuracy validation, and noise filtering. The processed data is then encapsulated into a LoRaWAN packet for transmission. The LoRaWAN module sends the packet over the LoRa network, which supports long-range communication (up to several kilometers) with low power consumption, to the ChirpStack server. The server decodes and stores the data in a database, such as InfluxDB, for further processing. Users can access and analyze the soil parameter trends through the Grafana interface, which visualizes data in charts, tables, or reports. The device is configured to collect measurements in 24-hour cycles, entering a low-power sleep mode after each measurement and transmission to conserve energy.

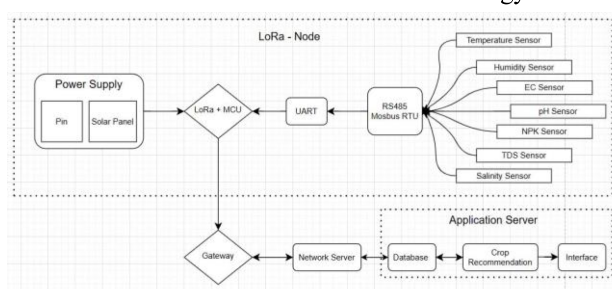

Figure 1. *Block diagram of system*

3.1. Hardware Design

The solar panel delivers a 6V input to the 18650 Li-ion battery charging circuit, which steps up the voltage to 12V to charge the 18650 batteries via a 3S 20A charging and protection module. The charged battery supplies 12V to a MOSFET module, which regulates power to the sensor based on control signals from the RF Thing UCA board. Additionally, the battery powers a 12V-to-5V conversion circuit, providing 5V to the UCA board and the TTL-to-RS485 module. The RS485 module decodes differential signals received on its A and B pins, converting them into TTL logic signals that are transmitted to the microcontroller’s RX pin. The microcontroller processes and normalizes the sensor data before formatting it for transmission via the LoRaWAN (Class A) protocol to the ChirpStack network server for storage and monitoring.

LoRaWAN defines three device classes: Class A, B, and C. Class A devices initiate uplink transmissions, followed by two short receive windows. Class B adds scheduled receive slots, while Class C keeps the receiver always on except during transmission. Due to regulatory requirements in Vietnam, transmissions follow a duty cycle limitation of 1% per hour. In real-world deployments, operational ranges can reach up to 2–3 km in urban areas and beyond 10 km in rural or open conditions.

To mitigate packet loss, the LoRaWAN system includes a retry mechanism with retransmission settings handled at the application layer. If a device does not receive an acknowledgment within 5 seconds, it resends the data. Timeout and retry limits can be adjusted based on deployment conditions.

Comparisons are made with similar research in Vietnam, such as a study by Phenikaa University, Hanoi, which also employed LoRa Class A devices and observed transmission ranges up to 2.1 km with retry mechanisms and performance metrics including confusion matrix and ROC curve.

To achieve this, we have developed and tested a prototype, focusing on wireless data transmission using the LoRaWAN protocol as shown in Figure 2. The LoRa module, connected to the microcontroller via TX and RX pins, is configured to transmit sensor data to a LoRa gateway, ensuring reliable data transfer for real-time soil monitoring [11].

Figure 2. *Soil quality measurement node*

3.2. Machine Learning Models for Recommendation Systems

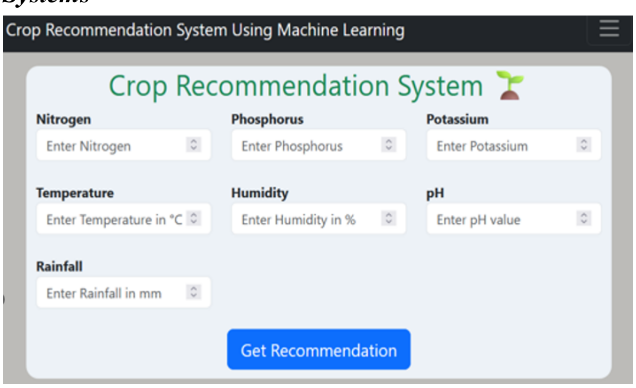

Figure 3. *Crop Recommendation System*

The machine learning model uses input data comprising temperature, moisture, pH, nitrogen (N), phosphorus (P), potassium (K), and rainfall measurements collected from a rainfall meter. The team tested various soil samples using an existing database to gather data on crop nutrient requirements. This data was analyzed by the machine learning model to derive actionable insights. The model's output is presented on the Crop Recommendation System web platform as shown in Figure 3, which recommends the most suitable crops for a specific area based on the measured soil quality parameters. The crop recommendation algorithm was developed using the Scikit-learn framework.

3.2.1. Databases

Several publicly available datasets on soil parameters were analyzed, including the "Crop Recommendation" dataset by Atharva Ingle [12], accessible on Kaggle. This dataset comprises 2,200 samples across 22 crops: rice, corn, chickpeas, red beans, peas, silk beans, green beans, black beans, lentils, pomegranate, banana, mango, grapevines, watermelon, muskmelon, apple, orange, papaya, coconut, cotton, jute, and coffee. Each sample corresponds to a specific crop and includes environmental factors such as temperature, soil pH, moisture, and rainfall, with the number of entries per crop varying based on these conditions.

The dataset features key parameters: nitrogen (N), phosphorus (P), and potassium (K) levels in the soil, average temperature, moisture, soil pH, and average rainfall. Designed for developing crop recommendation systems, it provides comprehensive data on crops, soil properties, and weather to predict the most suitable crop for a given area. This dataset supports classification tasks, enabling the creation of automated systems that recommend crops based on soil characteristics, climate, and environmental factors.

3.2.2. Machine Learning Models

The crop recommendation system is formulated as a classification problem, aiming to categorize each data sample into one of 22 crop classes based on the dataset. Machine learning models, such as Decision Trees and Random Forests, are well-suited for classification tasks, assigning each sample to a specific crop group. Various machine learning models were tested, with their performance evaluated to identify the most effective ones for crop recommendation.

Table 2. *Machine learnings model comparison*

Machine Learning Models	Accuracy
Logistic Regression	96.36%
GaussianNB	99.54%
SVC	96.81%
Kneighbors Classifier	95.9%
DecisionTree Classifier	98.4%
ExtraTree Classifier	86.13%
RandomForest Classifier	99.54%
Bagging Classifier	98.63%
GradientBoosting Classifier	98.18%

As shown in Table 2, various machine learning models were compared for the crop recommendation system, with the Random Forest Classifier selected due to its superior performance. Achieving an accuracy of up to 99.54%, Random Forest effectively handles non-linear data by partitioning the feature space into smaller regions for precise classifications. It robustly manages missing data, delivering reliable predictions despite incomplete inputs. By aggregating predictions from multiple decision trees, the model minimizes overfitting and enhances accuracy. These strengths make Random Forest well-suited for large datasets, such as the Crop Recommendation dataset with 2,200 samples, enabling accurate crop predictions based on soil and environmental parameters.

4. Results and discussions

4.1. Hardware

4.1.1. Energy Consumption

The device operates on a 12V, 2000mAh battery, supporting up to seven days of continuous operation without recharging. For improved efficiency and practicality, a solar battery charger was integrated to automatically recharge the battery. The solar panel provides a 6V input, which a booster circuit steps up to 12V, coupled with a battery protection circuit to prevent overcharging and prolong battery lifespan. To optimize power consumption, the processor's sleep mode feature was implemented, reducing energy use during periods when data packets are not transmitted to the Network Server. In sleep mode, the device minimizes current draw, thereby significantly extending its operational lifetime, as illustrated in Figure 4. Additionally, the MOSFET module cuts power to the sensor during sleep mode, further conserving energy and enhancing the device's longevity.

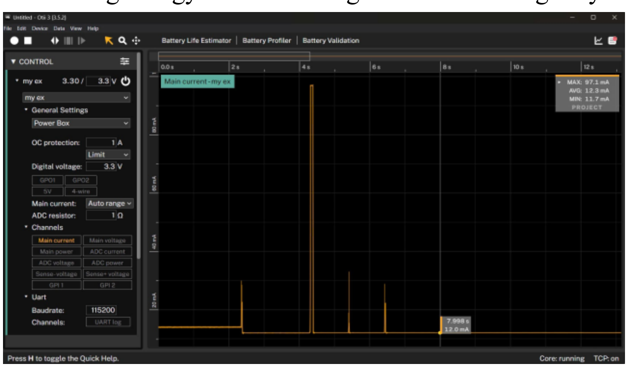

Figure 4. *Board UCA Current Measurement Ortic*

Table 3. *Power consumption in different mode*

Mode	Board current	Sensor current
Sleep Mode	12 mA	0 mA
Submit data	97.1 mA	17.8 mA

Based on the current consumption specifications of the device as presented in Table 3, it was calculated the device's operational time without the need to recharge the battery. With a total daily energy consumption of 288.28mAh and a battery capacity of 2000mAh, the device can operate continuously for 7 days before the battery is depleted.

4.1.2. Measured data from the sensor

Soil quality sensor accuracy was evaluated through measurements conducted at multiple locations. The integrated temperature-compensated sensor mitigates the influence of ambient temperature, ensuring precise readings. The Modbus RTU (RS485) protocol facilitates stable, interference-free signal transmission over long distances, as validated during tests at large-scale vegetable gardens, such as the Tuy Loan Vegetable Garden. Data transmitted to the Network Server accurately reflected the measurements, demonstrating high reliability without signal interference. Repeated measurements on identical soil samples produced consistent results, confirming the sensor's stable performance.

4.1.3. Ability to transmit information

Data transmission from the device occurs via LoRaWAN, with the following specifications: The transmission frequency ranges from 921.4 MHz to 922.8 MHz, selected to comply with regional frequency regulations. A Spread Factor of SF7 was chosen to balance fast data transfer with medium-range communication to the gateway, enabling efficient transmission over moderate distances. A standard bandwidth of 125 kHz (BW125) optimizes frequency band usage, enhancing signal quality.

Packet loss was assessed using the Frame Counter (FCnt) metric. Gaps in FCnt values, such as missing packets between FCnt:4061 and FCnt:4065, indicated packet loss. Despite these instances, the overall packet loss rate remains low. To improve signal reception over longer distances, the Spread Factor can be adjusted to SF10, which enhances sensitivity and supports remote signal acquisition. Table 4 compares the performance of SF7 and SF10 configurations.

Table 4. *Comparison of the results obtained using SF7 and SF10*

Number of transmissions	Spread Factor: SF7	Spread Factor: SF10
100 packets	94	100
100 packets (2nd time)	95	100
100 packets (3rd time)	95	99
100 packets (4 times)	93	99
100 packets (5 times)	94	99

4.2. Machine Learning Models

Figure 5. *Loss and Accuracy of training and testing data*

The Random Forest Classifier achieved the highest accuracy-up to 99.54%-among the nine machine learning models tested as illustrated in Figure 5. In the final training sessions, accuracy nearly reached 100%, with both loss and accuracy steadily improving and no signs of overfitting observed. These results are considered excellent for a machine learning model.

4.3. Practical experiments

After testing the model's effectiveness on a computer and obtaining measurement parameters from the device, a crop suggestion system was deployed on-site as shown in Figure 6. Nutrient data extracted from the soil was input into the system, which then captured an image and used the combined inputs to run predictions. This is one of several prototypes tested in the field. The system predicts the most suitable crop for the measured land based on soil quality data. However, other crops can still be grown on the same land, albeit with potentially lower suitability. The system's suggestions are based solely on the dataset used during training.

A key strength of the crop suggestion system, as shown in Figure 7, is its ability to deliver optimal results by considering various environmental factors, such as local weather conditions and rainfall. Since crop growth is highly dependent on weather, the system effectively collects and processes weather-related data to enhance prediction accuracy.

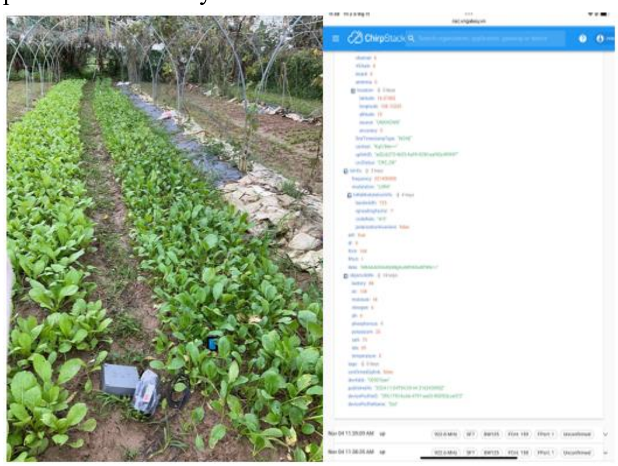

Figure 6. *On-site deployment and installation*

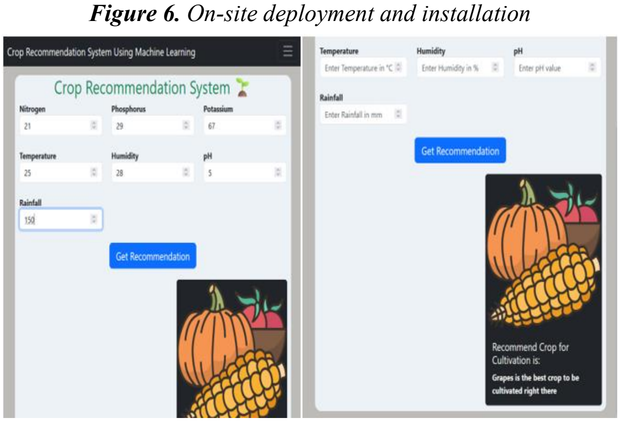

Figure 7. *Crop recommendations system interface*

5. Conclusion

In conclusion, the successful deployment and operation of the LoRaWAN-based system has proven to be an effective solution for soil monitoring and management in agriculture. The system has consistently collected key soil indicators, including moisture, temperature, EC, pH, and NPK content, while providing efficient, long-range data transmission with low energy consumption. The integration of machine learning algorithms presents significant potential for enhancing data analysis, predicting trends, and optimizing resource usage. Moreover, expanding the system to incorporate additional smart agricultural technologies, such as automated irrigation or crop management systems, will undoubtedly broaden its practical applications, offering a comprehensive and sustainable approach to modern agriculture.

Acknowledgement: This research is funded by The Murata Science Foundation and The University of Danang - University of Science and Technology, code number of Project: T2024-02-05MSF.

REFERENCES

[1] B. Loganathan, Ganesh, E. Mohan, and R. S. Kumar, "IoT based water and soil quality monitoring system", *International Journal of Mechanical Engineering and Technology (IJMET),* vol. 10, no. 2, pp 537-541, 2019.

[2] Bloem et al, "Monitoring and evaluating soil quality", *Microbiological methods for assessing soil quality Wallingford UK CABI Publishing*, pp 23-49, 2005.

[3] N. Ananthi, J. Divya, M. Divya, and V. Janani, "IoT based smart soil monitoring system for agricultural production", *IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR)*, 2017, pp. 209-214.

[4] P. Obade, Vincent, and R. Lal, "Assessing land cover and soil quality by remote sensing and geographical information systems (GIS)." *Catena,* vol. 104, pp 77-92, 2013.

[5] T. Hamrita and E. C. Hoffacker, "Development of a "smart" wireless soil monitoring sensor prototype using RFID technology", *Applied Engineering in Agriculture,* pp 139-143, 2005.

[6] U. K. Madhura, P. Akshay, A. J. Bhattad, and G. S. Nagaraja, "Soil Quality Management Using Wireless Sensor Network," *2017 2nd International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS)*, Bengaluru, India, 2017, pp. 1-5, doi: 10.1109/CSITSS.2017.8447860.

[7] T. Georgieva *et al*, "Design of wireless sensor network for monitoring of soil quality parameters", *Agriculture and Agricultural Science Procedia*, 2016, pp 431-437.

[8] Enfarm Agritech, "Enfarm soil nutrient measurement device", *enfarm.com* May 28, 2025. [Online]. Available: https://enfarm.com/product [Accessed June 03, 2025].

[9] Song long company, "Soil NPK meter zd2804", *slivn.com/*, July 1, 2020. [Online]. Available: https://slivn.com/may-do-npk-trong-dat-zd2804 [Accessed June 03, 2025].

[10] Hanna Instruments Vietnam, "Soil NPK and pH Measurement Kit", *hannavietnam.com/*, May 3, 2023. [Online]. Available: https://www.hannavietnam.com/detail-product/bo-dung-cu-do-npk-va-ph-cua-dat-25-lan-do-863 [Accessed June 03, 2025].

[11] M. Hamnache, R. Kacimi, and A. Beylot, "Enabling an inter-operator roaming capability in LoRaWAN networks", *Ad Hoc Networks*, vol 139, 2023.

[12] Kaggle, "Crop Recommendation", *kaggle.com*, Sep 12, 2020 [Online]. Available: https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset [Accessed June 03, 2025].