JRIIE Journal of Research Innovation and Implications in Education
Center for Research Implications & Practice

Website: www.jriiejournal.com ISSN 2520-7504 (Online) Vol.10, Iss.4, 2026 (pp. 2049 - 2057)

Development of a Decision Support System for Crop Selection Using A 7-in-1 Soil Sensor: A Case Study of Huye District, Rwanda

Mporanzi Joel and Musoni Willison
University of Kigali, Rwanda
Email: joelmporanzi9@gmail.com

Abstract: *This study developed and evaluated an Internet of Things (IoT)-based decision support system for crop selection using a 7-in-1 soil sensor in Huye District, Rwanda. The prototype combines a soil sensor and ESP32 microcontroller with a Flask web application, MySQL database and Random Forest machine learning model. The sensor measures nitrogen, phosphorus, potassium, pH, temperature, soil moisture and electrical conductivity, and transmits readings wirelessly for crop recommendation. A design science research approach and Agile development process were adopted. A total of 150 soil samples were collected from 50 farms in five sectors of Huye District. Because the field dataset was small and imbalanced across seven crop classes, model training and evaluation used a 2,100-record literature-grounded synthetic dataset, divided into training (1,470), validation (315) and test (315) subsets. Laboratory testing recorded a 95.3% first-attempt transmission success rate and an average response time of 27 seconds. On the synthetic held-out test set, Random Forest achieved 94.60% accuracy, 94.69% macro precision, 94.64% macro recall and a 94.57% macro F1-score, outperforming a Decision Tree baseline by 6.03 percentage points in accuracy. Forty participants evaluated usability, producing a mean System Usability Scale score of 76.9. The findings demonstrate the technical feasibility and promising usability of the prototype. However, the model's reported accuracy reflects performance on synthetic data and requires validation using a larger, independently labelled real-world soil dataset before operational deployment.*

Keywords: Decision support system; Internet of Things; soil sensor; crop recommendation; Random Forest; precision agriculture; Huye District; Rwanda.

How to cite this work (APA):

Mporanzi, J. & Musoni, W. (2026). Development of a Decision Support System for Crop Selection Using A 7-in-1 Soil Sensor: A Case Study of Huye District, Rwanda. *Journal of Research Innovation and Implications in Education, 10*(4), xxx – xxx. **https://doi.org/10.59765/cps2g**

1. Introduction

Agriculture remains central to livelihoods and economic activity in Rwanda. However, crop selection at farm level is often made without timely, plot-specific information about soil conditions. Farmers may rely on experience, informal advice or broad suitability information, even though soil properties can vary between neighbouring plots. Decisions made without adequate information about soil nutrients, acidity, moisture and temperature can contribute to inefficient use of seed and fertiliser and may expose farming households to avoidable production risks (Food and Agriculture Organization [FAO], 2023; Ministry of Agriculture and Animal Resources [MINAGRI], 2023).

Soil testing can provide information to guide crop selection, but laboratory testing may involve costs, travel and delays before results are returned. District-level soil suitability information is useful for planning, yet it may not reflect variation at the level of an individual farm. These limitations create a need for accessible tools that can capture soil conditions closer to where planting decisions are made and present the information in a form that farmers and agricultural extension officers can use.

Recent developments in the Internet of Things (IoT) have enabled sensors and connected devices to collect and transmit agricultural data. In precision agriculture, sensor-based monitoring can support more targeted management of soil and crop conditions (Mowla et al., 2023; Mansoor et al., 2025). The ESP32 microcontroller offers wireless connectivity for relatively compact IoT applications, while multi-parameter soil sensors can capture several indicators in one measurement process (Hercog et al., 2023; Prasad et al., 2024). Machine learning can then analyse measured soil characteristics and generate crop suitability predictions. Previous crop recommendation studies have reported strong performance from Random Forest models, although many rely on secondary datasets and manually entered soil values rather than integrated field sensing (Dey et al., 2024; Suresh et al., 2024).

This study addressed the need for an integrated, farm-oriented decision support system by developing a portable IoT prototype based on an ESP32 microcontroller and a 7-in-1 soil sensor, connected to a web dashboard and a Random Forest model. The study was guided by three objectives: (i) to design a portable IoT device capable of collecting and transmitting seven soil parameters; (ii) to develop a web application integrated with a Random Forest model for crop recommendation and storage of historical readings; and (iii) to test the integrated system and evaluate its predictive performance and usability among farmers and agricultural professionals in Huye District.

2. Literature Review

2.1 IoT and precision agriculture

Precision agriculture uses information about spatial and temporal variation in farm conditions to inform agricultural decisions. Soil sensors and connected monitoring systems can provide more frequent information than occasional manual observation or laboratory testing. Mowla et al. (2023) reviewed IoT and wireless sensor network applications in smart agriculture and identified soil moisture monitoring as a common application. Prasad et al. (2024) demonstrated the feasibility of real-time soil health monitoring using an ESP32-based system and multiple soil parameters. These studies support the technical case for sensor-based monitoring, while also indicating the importance of testing systems in relevant field contexts.

2.2 Machine learning for crop recommendation

Machine learning models can identify patterns between soil characteristics and crop suitability classes. Dey et al. (2024) and Suresh et al. (2024) reported high crop recommendation accuracy using Random Forest models on structured crop and soil datasets. Gajbhiye et al. (2025) similarly compared machine learning algorithms for crop recommendation and reported strong performance from ensemble approaches. Nevertheless, high predictive performance on a prepared dataset does not by itself demonstrate performance when measurements are collected by a physical sensor under variable field conditions. Integration of sensor acquisition, data transmission, prediction and user-facing explanation is therefore important.

2.3 Conceptual Framework of the Existing System

The existing crop-selection approaches share a common structure in which three independent inputs, farmer experience, RAB district-level suitability maps, and laboratory soil testing, feed into a manual, non-integrated interpretation process that produces a single dependent output: the farmer's planting decision. Because none of the three inputs is captured digitally or updated in real time, there is no feedback loop linking the outcome of a planting decision back into future soil measurement, and moderating factors such as farmer age, education, and weather variability are never systematically accounted for.

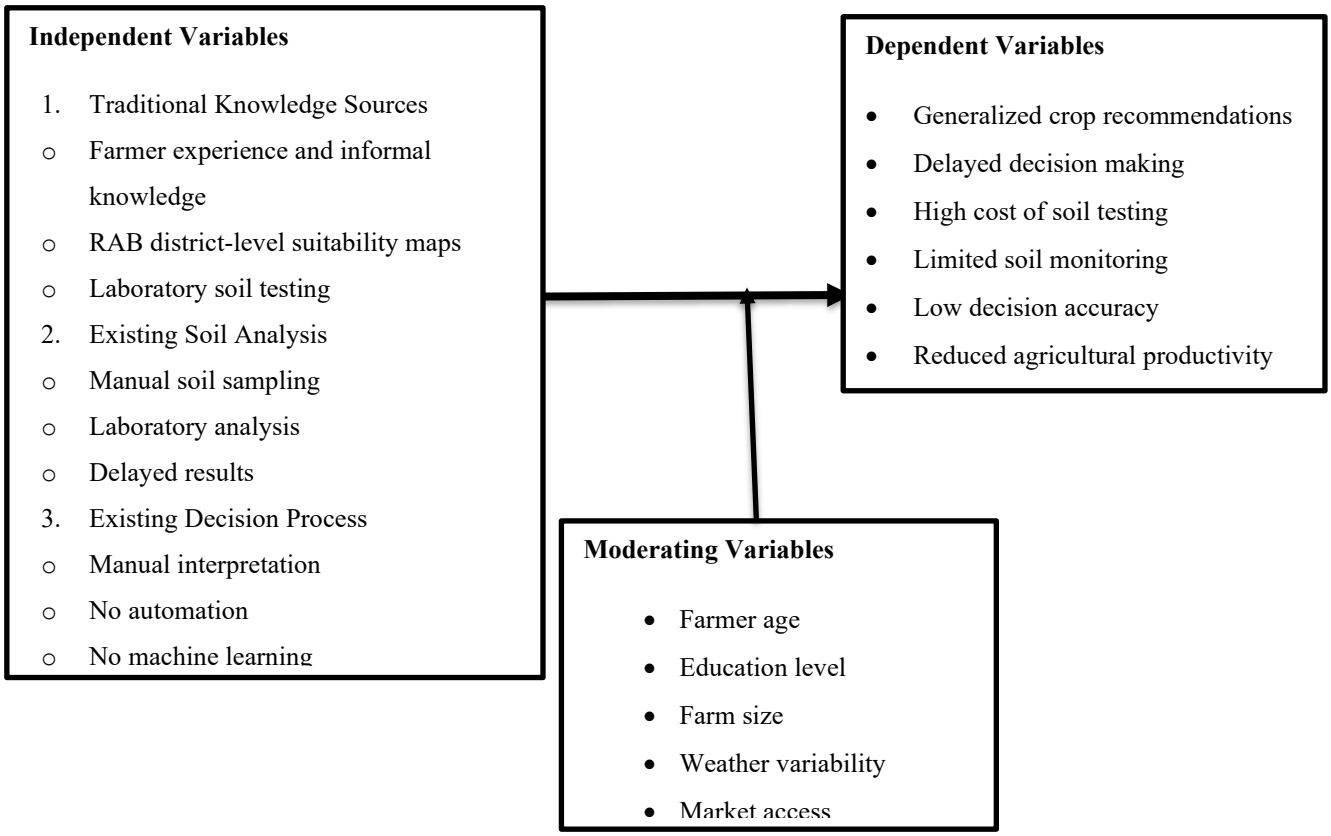

Figure 2.1: Conceptual Framework of the Existing System

2.4 Conceptual Framework of the Proposed System

Independent variables: the seven soil parameters the sensor reads from each farm (nitrogen, phosphorus, potassium, pH, temperature, soil moisture and electrical conductivity).

Process: three steps. An ESP32 reads and formats the sensor data, the data is sent over Wi-Fi to a web server, and a Random Forest model analyses it to produce a recommendation.

Dependent variables: the system's outputs, which are the crop recommendations plus displays of current soil readings and historical soil trends.

Moderating variables: farmer age, education level, farm size, seasonal weather and market access. The system does not measure these, but they are used as context when interpreting whether farmers follow a recommendation and how well it works.

Each component maps onto one specific objective. The soil parameters correspond to objective one (designing the IoT device). The process corresponds to objective two (building the web application and ML model). The outputs and their evaluation correspond to objective three (field implementation, accuracy and usability testing).

2.5 Research gap and contribution

The reviewed literature indicates progress in IoT-based soil monitoring and machine learning-based crop recommendation, but these functions are often considered separately. The present study combines a 7-in-1 sensor, ESP32 wireless transmission, a web dashboard and a Random Forest model in one prototype designed for smallholder farming conditions in Huye District. It also includes functional testing and a formal usability evaluation. Its contribution is a tested prototype and a documented local field dataset; it does not claim that the model has already been validated for large-scale agronomic deployment.

3. Methodology

3.1 Research design and development approach

The study adopted a design science research approach because its central purpose was to develop an artefact to address a practical problem and evaluate its performance. The artefact comprised the IoT hardware, web application, database and crop recommendation model. Agile software development was used to organise implementation into iterative stages covering hardware integration, backend and data pipeline development, machine learning and dashboard development, and system testing and refinement (Hevner & vom Brocke, 2023; Pargaonkar, 2023).

3.2 Study setting and soil sample collection

The study was conducted in Huye District, Southern Province, Rwanda, in five sectors: Ngoma, Huye, Karama, Simbi and Tumba. Ten farms were visited in each sector, giving 50 farms in total. Three soil samples were collected from different locations on each farm, resulting in 150 samples. Samples were taken from approximately 10–15 cm depth, placed in labelled bags and transported for laboratory testing. Before measurement, samples were prepared and moisture-conditioned using the procedure described in the dissertation.

3.3 IoT device and web application development

The hardware comprised a 7-in-1 soil sensor measuring nitrogen, phosphorus, potassium, pH, temperature, soil moisture and electrical conductivity; an ESP32 microcontroller; and a MAX485 RS485-to-TTL interface. The sensor communicated with the microcontroller using RS485 MODBUS. The ESP32 formatted the readings as a JSON payload and transmitted them over Wi-Fi to a Flask web application through an HTTP endpoint. The application stored readings and recommendations in a MySQL database. A React.js dashboard displayed current soil readings, crop recommendations and historical soil trends. The system also included local buffering to retain readings during simulated connectivity interruptions and synchronise them when connectivity returned.

3.4 Dataset preparation and machine learning

The study used two distinct datasets for different purposes. The field dataset consisted of 150 soil readings collected from 50 farms and was used for device testing and examination of local soil variation. The field labels were imbalanced across the seven crop categories, including 11 wheat and 8 coffee samples. The dissertation therefore reports that the field dataset was not used to train the classifier. Instead, a 2,100-record literature-grounded synthetic dataset was constructed using crop suitability ranges. The dataset was divided into 1,470 training records, 315 validation records and 315 test records. A Random Forest classifier was trained using the training subset, with model configuration selected using validation data. A Decision Tree classifier was trained on the same split as a baseline.

3.5 System and usability evaluation

Device performance was assessed using transmission success, response time and reading stability. Predictive performance was evaluated on the held-out synthetic test set using accuracy, macro precision, macro recall, macro F1-score and Cohen's kappa. The Random Forest model was compared with the Decision Tree baseline. System usability was evaluated with the System Usability Scale (SUS) among 40 participants comprising 20 smallholder farmers and 20 agricultural extension officers and agronomists. The SUS mean was compared with the benchmark score of 68. Qualitative feedback from potential users was grouped thematically. The dissertation reports that participation was voluntary and that confidentiality and anonymity were maintained.

4. Results and Discussion

The results are presented in relation to the study's three objectives: development and laboratory assessment of the IoT soil-sensing prototype; development and evaluation of the crop recommendation model; and functional and usability evaluation of the integrated system.

4.1 Field Soil Characteristics and IoT Device Performance

A total of 150 soil samples were collected from 50 farms across Ngoma, Huye, Karama, Simbi and Tumba sectors, with 30 samples obtained from each sector. The samples covered seven crop suitability categories: maize (25.3%), beans (21.3%), Irish potatoes (16.7%), cassava (14.0%), soybeans (10.0%), wheat (7.3%) and coffee (5.3%). The relatively small numbers for wheat and coffee, together with the overall sample size, limited the suitability of the field dataset for machine-learning training.

Table 1: Distribution of Crop Suitability Labels in the Field Dataset

Crop	Number of samples	Percentage
Maize	38	25.3%
Beans	32	21.3%
Irish potatoes	25	16.7%
Cassava	21	14.0%
Soybeans	15	10.0%
Wheat	11	7.3%
Coffee	8	5.3%

Note. The 150 field samples were distributed across seven crop suitability categories.

Maize and beans together accounted for 46.6% of the field labels, whereas wheat and coffee together represented 12.6%. This uneven distribution, particularly the small number of wheat and coffee samples, supports the decision to use a separate balanced synthetic dataset for model training.

Table 2: Descriptive Statistics of Field Soil Parameters (N = 150)

Soil parameter	Minimum	Maximum	*M*	*SD*
Nitrogen (ppm)	8.2	89.6	42.3	18.7
Phosphorus (ppm)	3.1	67.4	28.9	14.2
Potassium (ppm)	12.5	142.8	68.4	31.5
pH	4.3	7.2	5.8	0.7
Temperature (°C)	16.8	29.4	22.7	2.9
Soil moisture (%)	12.4	48.7	29.3	9.1
Electrical conductivity (dS/m)	0.12	1.87	0.64	0.38

Note. Values are based on the 150 field soil samples. M = mean; SD = standard deviation.

The measurements indicate variation in soil conditions across the sampled farms. Soil pH ranged from 4.3 to 7.2 (M = 5.8, SD = 0.7), showing that the samples included acidic through near-neutral conditions. Nitrogen, phosphorus and potassium also varied substantially, indicating that a single district-level soil profile would not represent all sampled farms.

The prototype combined the 7-in-1 sensor, ESP32 microcontroller and MAX485 RS485-to-TTL interface. During laboratory testing, 143 of 150 attempted readings were reported as successfully transmitted (95.3%). The average time from probe insertion to dashboard display was 27 seconds, while mean reading stability across the seven parameters was reported as a coefficient of variation of 2.3%. Seven readings were cached during simulated Wi-Fi interruptions and subsequently synchronised when connectivity returned.

Table 3: *Summary of IoT Prototype Laboratory Performance*

Performance measure	Result
Reading attempts	150
Reported successful transmissions	143 (95.3%)
Average time to dashboard display	27 seconds
Mean reading stability (CV)	2.3%
Offline readings cached and later synchronized	7

Note. CV = coefficient of variation. Testing was conducted under laboratory conditions; the results do not establish field performance under routine farm conditions.

4.2 Web Dashboard and Machine Learning Model

The web application integrated a Flask backend, React interface, MySQL database and Scikit-learn classification model. The 150 field readings were used to examine local soil variation and test the sensing and data pipeline. Because this dataset was small and imbalanced across the seven crop categories, the Random Forest model was trained and evaluated using a separate 2,100-record literature-grounded synthetic dataset. The records were divided into training (n = 1,470), validation (n = 315) and test (n = 315) subsets.

Table 4: *Feature Importance in the Random Forest Model*

Soil parameter	Importance	Rank
Nitrogen	27.9%	1
Phosphorus	20.3%	2
Potassium	14.9%	3
pH	11.8%	4
Temperature	11.2%	5
Electrical conductivity	7.2%	6
Soil moisture	6.8%	7

Note. Importance scores are the model's feature-importance estimates and sum to approximately 100%.

Soil moisture and electrical conductivity had lower relative importance, although all seven variables contributed to the model.

Nitrogen was the most influential model input (27.9%), followed by phosphorus (20.3%) and potassium (14.9%). Together, the three macronutrients accounted for 63.1% of the reported feature importance.

4.3 Predictive Performance

On the held-out synthetic test set, the Random Forest achieved 94.60% accuracy, compared with 88.57% for the Decision Tree baseline. The Random Forest also recorded higher macro-averaged precision, recall and F1-score. Its accuracy advantage was 6.03 percentage points.

Per-crop F1-scores ranged from 0.907 for beans to 0.989 for coffee, indicating that performance was not confined to a single crop category.

Table 5: Classification Performance on the Held-Out Synthetic Test Set

Metric	Random Forest	Decision Tree	Difference
Accuracy	94.60%	88.57%	+6.03 pp
Macro precision	94.69%	88.79%	+5.90 pp
Macro recall	94.64%	88.58%	+6.06 pp
Macro F1-score	94.57%	88.61%	+5.96 pp

Note. Each model was evaluated on the same test subset (n = 315). pp = percentage points. The Random Forest correctly classified 298 of 315 records.

Table 6: Per-Crop Classification Performance of the Random Forest Model

Crop	Precision	Recall	F1-score	*n*
Beans	0.975	0.848	0.907	46
Cassava	0.935	0.956	0.945	45
Coffee	0.978	1.000	0.989	45
Irish potatoes	0.977	0.956	0.966	45
Maize	0.918	1.000	0.957	45
Soybeans	0.911	0.911	0.911	45
Wheat	0.933	0.955	0.944	44
Macro average	0.947	0.946	0.946	315

Note. Precision, recall and F1-score are reported by crop class. n = number of test records in each class.

The validation accuracy was 95.24% for Random Forest and 88.89% for Decision Tree; the corresponding test accuracies were 94.60% and 88.57%. The small differences between validation and test results suggest similar performance across the two synthetic subsets. Beans had the lowest recall (0.848), with four bean records reported as misclassified as soybeans; two wheat records were classified as maize. Coffee and maize each had a recall of 1.000 in the test set.

The Random Forest exceeded the study's 80% accuracy benchmark (95% Wilson CI [91.5%, 96.6%]) and outperformed the Decision Tree by more than the prespecified five-percentage-point margin. The dissertation reports $p < .001$ for the test of the 80% benchmark and $p = .003$ for the model comparison. These results support the first and second hypotheses. However, because the evaluation data were synthetic rather than independently labelled field observations, the 94.60% result should be interpreted as performance on the synthetic test set, not as demonstrated predictive accuracy on real farms.

4.4 Functional Testing and User Feedback

All 10 functional test cases were reported as passed. The tests covered sensor data capture and validation, Wi-Fi transmission, offline caching and synchronisation, crop recommendation generation, historical trend display, MySQL storage, Flask API responses, dashboard rendering and role-based access. The system generated recommendations for the valid readings and rejected an invalid pH value during validation.

Feedback from farmers and agricultural professionals was generally positive. Participants valued the prospect of receiving soil information more quickly and considered the dashboard layout and colour-coded presentation helpful. They also recommended additional support, including a mobile application, SMS or voice-based access, clearer explanations of electrical conductivity, and fertiliser quantity guidance. These comments identify practical improvements for a future field-ready version.

4.5 Usability Evaluation

Usability was evaluated by 40 participants: 20 smallholder farmers and 20 extension officers/agronomists. The overall mean System Usability Scale (SUS) score was 76.9 (SD = 8.3), above the study benchmark of 68. Extension officers/agronomists reported a higher mean score (79.5) than farmers (74.2). In total, 34 participants (85%) scored above the benchmark, while six (15%) scored between 50.1 and 68.0.

Table 7: System Usability Scale Scores by Respondent Group

Respondent group	*n*	*M*	*SD*	Median	Range
Smallholder farmers	20	74.2	8.6	75.0	55–87.5
Extension officers	20	79.5	7.4	80.0	65–90
Overall	40	76.9	8.3	77.5	55–90

Note. SUS = System Usability Scale. M = mean; SD = standard deviation. Extension officers include agronomists. The study reports a one-sample t-test against 68, $t(39) = 6.78$, $p < .001$.

The usability result supports the third hypothesis that the prototype would achieve a mean SUS score above 68. The difference between professional and farmer scores suggests that the interface may require further adaptation for users with varying levels of digital familiarity. Overall, the findings indicate promising perceived usability, while further testing in routine farm settings is needed.

4.6 Discussion of Findings

This study developed and evaluated an integrated decision support prototype that combines a 7-in-1 soil sensor, an ESP32 microcontroller, wireless data transmission, a web dashboard and a Random Forest crop recommendation model. The findings are discussed in relation to the three study objectives and the literature reviewed. Overall, the results demonstrate the technical feasibility and promising perceived usability of the prototype under the conditions tested. However, the predictive results must be interpreted cautiously because the classifier was trained and evaluated using synthetic rather than independently labelled field data.

4.6.1 Soil variation and the value of farm-level sensing

The 150 soil samples collected from 50 farms across five sectors showed variation in nitrogen, phosphorus, potassium, pH, temperature, moisture and electrical conductivity. For example, pH ranged from 4.3 to 7.2, while potassium ranged from 12.5 to 142.8 ppm. This spread indicates that soil conditions differed across the sampled farms and supports the rationale for collecting measurements at farm level rather than relying exclusively on a general district profile. The finding is consistent with the precision-agriculture perspective that spatial variation in soil and crop conditions should inform management decisions (Mansoor et al., 2025; Sishodia et al., 2020).

The result also aligns with Mowla et al. (2023), who identified sensor networks as useful tools for monitoring agricultural conditions, and with Prasad et al. (2024), whose ESP32-based system demonstrated the feasibility of collecting multiple soil-health indicators. The present prototype extends this approach by connecting sensor readings to a crop recommendation workflow. Nevertheless, the observed variation in the field dataset does not, by itself, establish which crop will achieve the highest yield on each farm. Crop performance also depends on rainfall, management practices, seed variety, pests, market access and other conditions that were outside the predictive model. Farm-level sensing should therefore be understood as improving the information available for decisions, not as replacing agronomic judgement.

4.6.2 Technical performance of the IoT prototype

The prototype transmitted 143 of 150 readings successfully on the first attempt (95.3%), displayed readings after an average of 27 seconds and recorded a mean reading-stability coefficient of variation of 2.3%. The seven readings affected by simulated connectivity interruptions were cached and later synchronised. In addition, all ten functional test cases passed. Taken together, these results indicate that the hardware, communication interface, server and dashboard operated as an integrated system in the controlled laboratory environment.

These findings are broadly consistent with the technical feasibility reported in studies of ESP32-based IoT devices and multi-parameter agricultural monitoring (Hercog et al., 2023; Prasad et al., 2024). They also respond to the wider potential of IoT and wireless sensor networks for smart agriculture described by Mowla et al. (2023). The offline caching function is particularly relevant to settings where connectivity may be intermittent. However, laboratory performance should not be equated with reliability during routine farm use. The prototype still requires longer-duration field testing under changing weather, power and network conditions. In addition, the sensor's readings should be calibrated and compared with laboratory reference measurements before the system is used for consequential agronomic recommendations.

4.6.3 Interpretation of the crop recommendation model

On the held-out synthetic test set, Random Forest achieved 94.60% accuracy, compared with 88.57% for the Decision Tree baseline. The Random Forest also recorded higher macro-precision, macro-recall and macro-F1 scores. Its 6.03-percentage-point accuracy advantage exceeded the study's prespecified five-point margin. This result is consistent with the reviewed crop recommendation studies in which Random Forest performed strongly relative to alternative classifiers (Dey et al., 2024; Gajbhiye et al., 2025; Suresh et al., 2024). The ensemble structure of Random Forest can reduce reliance on the decisions of a single tree and may help capture interactions among soil variables.

The feature-importance results showed that nitrogen, phosphorus and potassium together accounted for 63.1% of the model's reported importance, with pH, temperature, electrical conductivity and soil moisture providing additional information. This pattern is agronomically plausible because nutrient availability and soil acidity are important considerations in crop suitability. However, feature-importance scores describe how the fitted model used the input variables; they do not establish that a variable independently causes crop success, nor do they quantify the effect of changing that variable in the field. Their interpretation should remain within the limits of the model and dataset.

Per-crop results were generally high, with F1-scores ranging from 0.907 for beans to 0.989 for coffee. Beans had the lowest recall (0.848), and the dissertation's confusion-matrix interpretation indicates that some bean records were classified as soybeans. This pattern may reflect overlapping soil suitability ranges represented in the synthetic data. It is useful because it identifies a class distinction that warrants further examination, but it should not be taken as evidence that the same confusion will occur at the same rate on actual farms.

A central qualification is that the 2,100 records used to train and test the classifier were literature-grounded synthetic records. The 150 field samples were used to examine local soil variation and test the sensing pipeline; they were not used as the independent test set for the

reported model accuracy. Consequently, the 94.60% accuracy indicates how well the model classified records generated from the encoded suitability ranges. It does not demonstrate 94.60% accuracy on real-world soil samples from Huye District. The higher accuracies reported by Dey et al. (2024) and Suresh et al. (2024) should likewise not be compared directly without accounting for differences in dataset construction, input variables, class definitions and validation procedures. Independent testing on a larger, labelled field dataset is necessary before operational use.

4.6.4 Usability and the prospects for farmer adoption

The overall System Usability Scale score was 76.9 (SD = 8.3), above the study benchmark of 68. The mean score among extension officers and agronomists (79.5) was higher than that of smallholder farmers (74.2), while 34 of the 40 participants (85%) scored above the benchmark. These findings suggest that participants generally considered the prototype usable after demonstration. The result is comparable with the positive usability assessment reported for an agricultural decision-support application by Karunathilake et al. (2023), although differences in systems, participants and evaluation settings mean that the scores should be treated as contextual rather than direct comparisons.

The slightly lower average among farmers suggests that further user-centred refinement may be needed for people with different levels of digital experience. This interpretation is consistent with McCampbell et al. (2023), who emphasised the importance of perceived usefulness and ease of use in the readiness of Rwandan farmers to use phone-based agronomic advice. It also reflects the wider adoption challenges identified in East African digital agriculture, including connectivity, affordability, training and digital literacy (Foster et al., 2023; Mhlanga, 2023; Mudereri et al., 2024). Participants' requests for mobile access, SMS or voice support, local-language explanations and clearer interpretation of electrical conductivity provide practical priorities for future development.

A favourable SUS score should not, however, be interpreted as proof of sustained adoption or improved agricultural outcomes. Participants evaluated a prototype in a study setting; actual adoption will also depend on cost, access to compatible devices, trust in recommendations, extension support and whether subsequent farm experience confirms the advice. Longer-term field pilots are therefore needed to examine continued use and the effect of the system on crop-selection decisions.

4.6.5 Overall contribution and implications

The principal contribution of the study is the integration of soil sensing, wireless transmission, data storage, machine learning and a user-facing dashboard in a prototype designed around the needs of smallholder agriculture in Huye District. The study also documents soil measurements from 50 farms and evaluates the system from technical and user perspectives. This responds to the gap identified in the literature between IoT systems that monitor soil conditions and crop recommendation models that commonly rely on manually entered or secondary data (Dey et al., 2024; Mowla et al., 2023; Prasad et al., 2024).

The evidence supports describing the prototype as technically feasible under laboratory conditions and promising in perceived usability. It does not yet support claims of validated field-level prediction accuracy, increased yields, reduced production costs or improved farmer income. The next research stage should prioritise sensor calibration against laboratory tests, collection of a larger and more representative labelled field dataset across seasons, independent model validation, and field trials involving farmers and extension officers. This staged approach would provide the evidence needed to determine whether the prototype can move from a working research artefact to a dependable agricultural decision-support service.

5. Conclusion and Recommendations

5.1 Conclusion

This study developed and evaluated an IoT and machine learning-based decision support system for crop selection using a 7-in-1 soil sensor in Huye District, Rwanda. The prototype integrated an ESP32 microcontroller, a soil sensor, a Flask application, a MySQL database, a React dashboard and a Random Forest classifier. Laboratory testing recorded a 95.3% first-attempt transmission success rate and an average response time of 27 seconds. The Random Forest model achieved 94.60% accuracy on a held-out synthetic test set and outperformed the Decision Tree baseline. The overall SUS score of 76.9 indicated above-benchmark perceived usability among the 40 participants.

These findings demonstrate the technical feasibility and promising usability of the prototype. They do not yet establish the model's predictive accuracy under real farm conditions, because the model was trained and evaluated using synthetic data. The system should therefore be treated as a prototype requiring further validation rather than as a replacement for laboratory soil analysis or professional agronomic advice.

5.2 Recommendations

Based on the findings of this study, the following recommendations were made:

1. The model should be retrained and independently validated using a substantially larger dataset of real soil measurements and crop suitability labels collected across multiple seasons and locations.
2. The sensor should be calibrated and its readings compared systematically with laboratory reference measurements.
3. The prototype should be piloted on working farms under field conditions, including variable weather, power supply and network connectivity.
4. The dashboard should be refined in response to user feedback by adding mobile access, SMS or voice-based interaction, Kinyarwanda support and plain-language explanations of soil parameters.

5. Agricultural institutions may examine the potential for linking farm-level sensing with existing soil information and extension services, subject to further technical validation.

References

Dey, B., Abir, K. A. M., Ahmed, R., Salam, M. A., Redowan, M., Miah, M. D., & Iqbal, M. A. (2024). Crop recommendation system using machine learning. *ITEGAM Journal of Engineering and Technology for Industrial Applications, 10*(48), 63–68. https://doi.org/10.5935/jetia.v10i48.884

Food and Agriculture Organization. (2023). *The state of food and agriculture 2023: Revealing the true cost of food*. FAO.

Foster, L., McCampbell, M., Sumberg, J., & Wellard, K. (2023). Smart farming and artificial intelligence in East Africa: Addressing indigeneity, plants, and gender. *Smart Agricultural Technology, 3*, Article 100132. https://doi.org/10.1016/j.atech.2022.100132

Gajbhiye, M., Kuhite, R., Farkade, V., Choudhary, S., & Dhomne, S. (2025). Crop recommendation using machine learning. *International Journal on Advanced Electrical and Computer Engineering, 14*(1), 138–144. https://journals.mriindia.com/index.php/ijaece/article/view/402

Hercog, D., Lerher, T., Truntic, M., & Tezak, O. (2023). Design and implementation of ESP32-based IoT devices. *Sensors, 23*(15), Article 6739. https://doi.org/10.3390/s23156739

Karunathilake, E. M., Le, A. T., Heo, S., Chung, Y. S., & Mansoor, S. (2023). The path to smart farming: Innovations and opportunities in precision agriculture. *Agriculture, 13*(8), Article 1593. https://doi.org/10.3390/agriculture13081593

Mansoor, S., Iqbal, S., Popescu, S. M., Kim, S. L., Chung, Y. S., & Baek, J. H. (2025). Integration of smart sensors and IoT in precision agriculture: Trends, challenges and future prospectives. *Frontiers in Plant Science, 16*, Article 1587869. https://doi.org/10.3389/fpls.2025.1587869

McCampbell, M., Sumberg, J., Wellard, K., & Foster, L. (2023). Are farmers ready to use phone-based digital tools for agronomic advice? Ex-ante user readiness assessment using the case of Rwandan banana farmers. *The Journal of Agricultural Education and Extension, 29*(1), 29–51. https://doi.org/10.1080/1389224X.2021.1984955

Mhlanga, D. (2023). Digital technology adoption in the agriculture sector: Challenges and complexities in Africa. *Human Behavior and Emerging Technologies, 2023*, Article 6951879. https://doi.org/10.1155/2023/6951879

Ministry of Agriculture and Animal Resources. (2023). *Rwanda agriculture sector strategic plan 2023–2026*. MINAGRI.

Mowla, M. N., Mowla, N., Shah, A. F. M. S., Rabie, K. M., & Shongwe, T. (2023). Internet of Things and wireless sensor networks for smart agriculture applications: A survey. *IEEE Access, 11*, 145813–145852. https://doi.org/10.1109/ACCESS.2023.3346299

Mudereri, B., Gacheru, S., Nkurunziza, L., & Ndegwa, M. (2024). *Rwanda–Smart Nkunganire System (SNS) fertilizer recommendation tool: Innovation packaging and scaling readiness (IPSR) analysis for smart agronomy*. International Potato Center. https://doi.org/10.4160/cip.2024.07.006

Prasad, S., Desai, N., & Kumar, A. (2024). IoT-based real-time soil health monitoring system for precision agriculture. *International Research Journal of Engineering and Technology, 11*(7), 470–478.

Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of remote sensing in precision agriculture: A review. *Remote Sensing, 12*(19), Article 3136. https://doi.org/10.3390/rs12193136

Suresh, G., Kumar, D. A. S., Lekashri, D. S., & Manikandan, D. R. (2024). Enhancing agricultural productivity: A machine learning approach to crop recommendations. *Human-Centric Intelligent Systems, 4*(2), 1–14. https://doi.org/10.1007/s44230-024-00081-3