

An Intelligent Crop Lifecycle Automation and Soil NPK Estimation System Using TFLite and Sensor Fusion

^[1] Piyush Gawali, ^[2] Shrirang Amle, ^[3] Rohit Raut, ^[4] Harsh Mishra, ^[5] Shriram Vyavahare

^{[1][2][3][4][5]} Department of Computer Engineering, Vishwakarma Institute of Technology Pune,

Affiliated with Savitri Bai Phule Pune University (SPPU), India

Email: ^[1] piyush.gawali@vit.edu, ^[2] shrirang.amle24@vit.edu, ^[3] rohit.raut24@vit.edu, ^[4] harsh.mishra24@vit.edu, ^[5] shriram.vyavahare24@vit.edu

Abstract— The trend of pressure on the controlled crop farming and sustainable agriculture has led to the growing urgency of the search of the reliable indoor crop production and economical means of soil health analysis. Traditional indoor farming technologies tend to be based on manual control, fixed control measures, or costly laboratory-based soil analyzes, which is why they are inaccessible and inefficient in continuous terms. This paper introduces a multi-purpose smart indoor crop growing and on-field soil intelligence system that is aimed at offering automated environmental regulation, real-time monitoring, and low-cost soil nutrient estimation. To control temperature, humidity, and irrigation, the indoor cultivation unit is controlled by a deterministic, rule-based control structure that uses a four-stage crop life-cycle model to utilize real-time sensor feedback. The two-layered safety system allows quicker detection of anomalies and continued recording of data that will be used to tolerate faults. An on-field soil analysis module to aid cultivation decisions is indirectly estimated through low-cost sensors and smartphone-based imaging with lightweight and edge-deployable machine learning models to measure soil pH, type, and NPK. The system is completely offline and it does not rely on lab test or specialized NPK sensors. In experimental analysis, the system exhibits constant environmental control, consistent alerts, and correct inference of soil nutrients, which provides a low-cost and scalable solution to precise indoor agriculture and sustainable crop management.

Index Terms— Indirect NPK Estimation, Sensor Fusion, Edge AI, Soil NPK Prediction, Crop Life-Cycle Control, Serial Communication, Fault Tolerance, Real-Time Monitoring

I. INTRODUCTION

The issues that are threatening Agriculture include climate variability, soil degradation, ineffective exploitation of resources and the escalating cost of conventional soil testing and crop management techniques. Proper understanding of the nutrients in the soil, the environmental factors and the stage of crop growth is very vital in ensuring sustainable production and the production of regular harvests. Nevertheless, the traditional soil analysis techniques that are conducted in the laboratory are usually time consuming, costly, and not affordable to small and marginal farmers especially in the rural areas. Also, the current state of manually tracking the environment of crops, particularly in a controlled or indoor farming system, is labor-intensive, subject to human error, and cannot support real-time decision making. Consequently, there is the increased demand of smart, automated and economical farming solutions that can perform consistently without significant reliance on external infrastructures.

Recent developments in precision agriculture have brought sensor based monitoring and automation to enhance efficiency in farming and minimize the need to use hands. Although these systems do offer real-time environmental information, most of them rely on fixed rule-based control systems or cloud-dependent analytics, which restricts their flexibility, scalability and reliability in resource-limited and low-connectivity environments. Moreover, the majority of

the current soil nutrient estimation methods rely on direct NPK sensors or those that were produced in the laboratory, which is considerably expensive and impossible to scale up to a large scale on-field use by small and marginal farmers. The above restrictions underscore the necessity of integrated, inexpensive, and locally exploitable solutions capable of providing actionable insights without having to rely on extensive infrastructure support.

To address these constraints, this study suggests a combined agricultural intelligence system, which is a combination of automated management of the indoor crop environment and the low-cost on-field soil quality analysis. The indoor cultivation module is created to maintain the essential growth conditions constantly across the whole crop lifecycle where crops grow in stable and controlled conditions with minimum human involvement. The system facilitates the same by adjusting the operating conditions in different growth phases, which ensures the monitoring consistency, timely response to deviations, and crop safety in the long-term cultivation. Simultaneously, the on-field module goes beyond the controlled setting of decision-making, offering actionable insights into soil health in the farm level. The system uses the available sensing and image-based analysis to determine the nutrient features of soil, instead of using laboratory testing or costly hardware. The combination of the two complementary parts creates a single solution that will bridge the controlled growth and real-world soil intelligence to make informed agricultural decisions that

are cost-effective, scalable, and applicable across different farming environments.

The system proposed focuses on offline functioning, modularity, and minimal computational cost, which makes it a reliable system to be used in controlled indoor and open-field agricultural applications. The framework will provide an accurate decision-making process by integrating the automation of the environment with the artificial intelligence of the soil, which will minimize operational expenses and encourage sustainable agriculture using accessible, scalable technology.

II. LITERATURE REVIEW

In V. S. Patyal et al. (2023), the authors determined the operational efficiency of 48 Indian electricity distribution utilities through a hybrid approach of DEA-IRP-TOPSIS. The analysis has found that there are considerable differences in efficiency between utilities and highlighted the best and the worst performers that can be used to implement effective measures to enhance efficiency in power distribution. [1]

G. Kansal and R. Tiwari (2024) suggested an incentive-based demand response framework based on price, to fit the modern Indian distribution systems. The method combines an augmented incentive-based demand response (IBDR) scheme that is effective in increasing the load flexibility and decreasing the peak demand, as confirmed by simulating on real grid models. [2]

S. Ghosh et al. (2023) introduced a neural network based fault location algorithm to 11 kV Indian power distribution lines with high accuracy and fast fault detection, which is useful in enhancing the reliability and response time of the system in case of a fault. The model uses the voltage and current measurements to localize the faults accurately. [3]

K. Singh and N. Mudgal (2024) compared how the power distribution companies of Haryana evolved into profit making units. They gave benchmarks on operational reforms and financial performance gains, which can be extended to other Indian DISCOMs to become financially sustainable. [4]

G. Gupta et al. (2025) conducted a review of modernization of the electricity market of India and paid attention to the market-based power purchase agreements (PPAs), the contracts for difference and the revenue-sharing contracts. The paper explains regulatory issues and suggests enabling processes that will create a competitive, flexible, and sustainable electricity market. [5]

In their research, S. K. Sahoo et al. (2024) evaluated the impact of hydroelectric generation at the level of increasing the resilience of the Indian power grid in the case of extraordinary events. The authors note the flexibility and quick response of hydro as crucial in the stability of the grid and balancing intermittent renewable reduction [6]

D. Dubey et al. (2024) suggested a framework of efficiency evaluation based on the integration of DEA and machine learning approaches to benchmark the Indian electricity distribution companies. The model is better at predicting the future and gives strong performance evaluations to inform policy and operation decisions. [7]

A. Kumar and M. Singh (2023) have created a deep learning-based fault location model of Indian power distribution networks based on low voltage measurements data. The method enhances the speed and accuracy of fault detection, therefore, allowing timely correction and reduced outage periods. [8]

B. O. Olorunfemi et al. (2023) performed an elaborate review of the multi-agent system implementation in the demand response management aspect of Indian energy distribution. They determined possible load optimization methods and pointed out the advantages of decentralized control structures to optimize energy consumption. [9]

The paper shows that the data envelopment analysis to test the technical efficiency of 55 electric distribution utilities in India. The research establishes essential performance lapses and suggests ways of enhancing efficiency of the entire systems and quality of customer service. [10]

III. METHODOLOGY

The suggested system combines automated control of the indoor crop environment with an autonomous on-field soil intelligence module to aid precision agriculture with the help of a hybrid cyber-physical architecture. The methodology is divided into two modules (A) Indoor Crop Lifecycle Automation Module and (B) On-Field Soil Intelligence Module. The two modules are independent of each other and yet they all combine to make informed decisions in agriculture.

3.1 Indoor Crop Cultivation and Monitoring Module

The module guarantees the optimum growth of crops in an indoor setup through a control of the environmental conditions that is constantly regulated with the help of rule-based control, real-time sensing and fault-tolerant monitoring.

3.1.1 System Architecture and Communication

A Master-Slave architecture is followed, in which a Python-based supervisory application runs high-level control code and user interaction, and an Arduino Nano is used to get sensor readings and provide actuator control. Connection is made through non-blocking USB serial connection. Control commands are event-driven and sensor telemetry is periodically reported to make sure that behavior is deterministic (not random) and that the serial buffer does not become overrun. It uses a lightweight delimiter based protocol e.g. 'CropCycle:120, DayCount:33, CurrentTemp:24.5, Current Humidity:60.0, SoilTemp:22.1, SoilMoisture:45.0, ...'

This protocol is a delimiter-based protocol with low parsing overhead and no saturation of serial buffers.

3.1.2 Environmental Data Acquisition and Filtering

Analog sensors that are connected to the Arduino Nano are used to monitor air temperature, soil temperature, and soil moisture. In order to reduce electrical noise and sporadic fluctuations, moving average filter is used:

$$V_{stable} = \frac{1}{N} \sum_{i=1}^N V_{raw,i}$$

where $N=10$ samples/acquisition cycle. Control decisions are made only using the filtered values.

3.1.3 Adaptive Crop Lifecycle Control Algorithm

The system follows a predefined four-stage crop lifecycle. The total growth duration T_{total} is divided into equal stages:

$$T_{stage} = \frac{T_{total}}{4}$$

The active growth stage $S_{current}$ is computed as:

$$S_{current} = \left\lfloor \frac{D_{current}}{T_{stage}} \right\rfloor$$

Based on $S_{current}$, temperature and humidity set points are automatically updated. The hysteresis logic is used in actuator control to avoid the frequent switching and provide stability to the system.

3.1.4 Safety and Alarm Mechanisms:

Two layers of safety framework is used to increase the reliability:

- 1. Rapid Change Detection:** First-order difference analysis is used to detect abnormal sensor spikes:

$$\Delta X = |X_{current} - X_{previous}|$$

When the change in X is more than a predetermined value during a short period, an emergency alarm is raised.

- 2. Setpoint Deviation Detection:** A latching system determines long-term deviation of target conditions, and the alarms are not shut off until some corrective action has been performed.

To ensure robustness during power interruptions:

- EEPROM state saving makes it possible to save parameters like crop day count and environmental setpoints which are important.
- Non-blocking millis()-based scheduling does away with blocking delays and improves responsiveness.

3.1.5 Data Logging and Fault Tolerance:

The system will allow the administrative team to log data and withstand fault conditions to prevent data loss in case of a failure. All sensor values and control conditions are recorded on daily CSV files to be able to trace and analyze them later. Manual intervention is controlled by a state machine (Start Active Stop) which provides safe state changes between automated and manual mode.

3.2 On-Field Soil Intelligence Module:

The module will be designed to allow estimating the intelligence of soil on the field using a sensor to measure the soil composition, moisture, temperature, and salinity. This module facilitates the ability to measure the quality of the soil and estimate the nutrient directly in the field at low costs through indirect sensing and edge-based machine learning.

3.2.1 Soil Data Acquisition:

Low-cost sensors are used to measure physical parameters (soil moisture and temperature). Moreover, the images of soil surface and pH strips are taken with the help of a smartphone camera according to the standard soil-water preparation.

3.2.2 Image Processing Pipeline

The images obtained are transformed into the HSV color space to increase the resistance to the changes in the illumination. Strong values of hues are obtained at a chosen area of interest and compared with some calibrated reference values. Color similarity is calculated with the help of Delta-E:

$$\Delta E = \sqrt{(L_2 - L_1)^2 + (a_2 - a_1)^2 + (b_2 - b_1)^2}$$

The nearest is what defines the soil color category and the pH.

3.2.3 Feature Vector Construction and Normalization

All the parameters obtained are merged into one feature vector:

$$X = [\text{Moisture}, \text{Temperature}, \text{Humidity}, \text{Soil Type}, \text{Soil Color}, \text{pH}]$$

Standardization features This is done by normalizing features (Z-score) to ensure similar inference across devices:

$$X_{norm} = \frac{X - \mu}{\sigma}$$

Where, μ and σ are derived from the training dataset.

3.2.4 NPK Prediction and Crop Recommendation

The converted Nitrogen (N), Phosphorus (P), and Potassium (K) concentrations of the soil are predicted using a Random Forest regression model converted to TensorFlow Lite. Another classification model produces crop and fertilizer prescriptions using a forecasted nutrient and soil property. Every inference is done on the mobile device and this allows full offline functionality.

IV. SYSTEM ARCHITECTURE

The proposed system is developed as a hybrid cyber-physical architecture system combining automated indoor crop management with an autonomous on-field soil intelligence system. The architecture is of a modular nature giving it flexibility in deployment, scalability and fault isolation. Logically, the system consists of two major subsystems, which include Indoor Crop Lifecycle Automation Subsystem and On- Field Soil Intelligence Subsystem, which act independently but aid in making precise agricultural decisions.

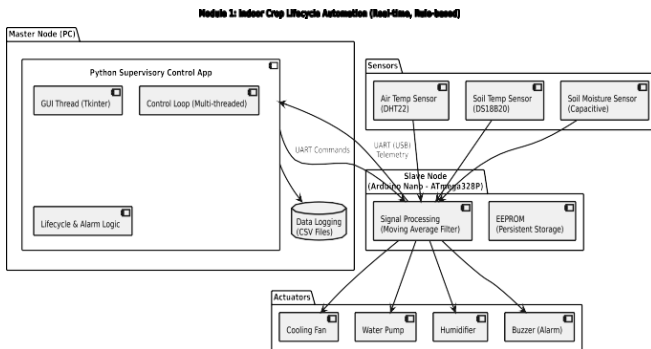


Fig 1. Indoor Crop Lifecycle Automation Subsystem

The indoor automation system has the role of ensuring that the environment in controlled agricultural areas like polyhouses or indoor farms maintains optimal conditions that support the growth of crops. It is made up of a Python based supervisory unit (PC) and an Arduino Nano based hardware controller, which are connected through a USB serial interface.

The supervisory unit is the pivotal decision-making unit that regulates the lifecycle of the crops, environmental parameters, user interaction and data recording. Arduino Nano is constantly receiving sensor data, pre-filtering it, and carrying out actuator commands sent to it by the supervisory system. Non-blocking communication protocol can be used to guarantee a reliable exchange of data without delay or buffer congestion.

This subsystem facilitates real time monitoring and automatic activation of the environmental parts such as cooling fans, irrigation pumps, and humidifiers. It also includes safety measures and non volatile memory storage in order to maintain continuous functionality in the event of power outages or any other unexpected system reboots.

On-field soil intelligence subsystem is the technology that allows performing the soil analysis at the field level and estimating nutrients at reasonable costs. It consists of simple soil sensors connected to a microcontroller and a processing unit that is based on the smartphone. The soil moisture and temperature are registered, and the soil surface photos and the pH strip photos are taken with the help of the smartphone camera.

All the processing and inference of data are done locally on the mobile device through edge-based machine learning models. Such design removes reliance on cloud infrastructure, which makes the system appropriate to be deployed in rural and connectivity-limited settings. The subsystem offers classification of the soil type, estimation of the pH, prediction of the NPK and recommendations of crops using indirect sensing methods.

4.1 System Integration and Deployment Strategy:

The two subsystems are conceptually aligned in a coherent precision agriculture framework in spite of the fact that they work independently. The subsystem of indoor automation is aimed at constant control over the environment, whereas the subsystem of on-field gives regular information about the health of the soil. This isolation provides a modularity, less complexity in the system and provides the ability to scale or deploy systems independently as per the needs of agriculture.

The suggested architecture guarantees efficient cost implementation, offline stability, and scalability to various agricultural conditions, which is appropriate to real-world precision farming applications, both small and large-scale indoor farms, as well as open field farming.

V. RESULTS

5.1 Indoor Crop Lifecycle Automation Performance:

The indoor crop cultivation module was tested in continuous mode to test whether it could uphold predetermined temperature and humidity setpoints at various lifecycle phases of crops. The system was able to control environmental conditions based on real-time sensor feedback and rule-based hysteresis control.

The results observed included stable actuator behavior where the actuator did not chatter the relay even in fast changing environmental conditions. Findings show that the temperature was kept at a range of +1.2 degrees centigrade of the target set point and the deviation of humidity was within +4%. Timely actuation was guaranteed by the non-blocking control loop and automatic stage transition allowed the process of crop growth progression to take place without any human intervention. Measurement: Accuracy of Crop Lifecycle Automation: CLA results are evaluated by determining the accuracy of CLA in crop lifecycle stage prediction.

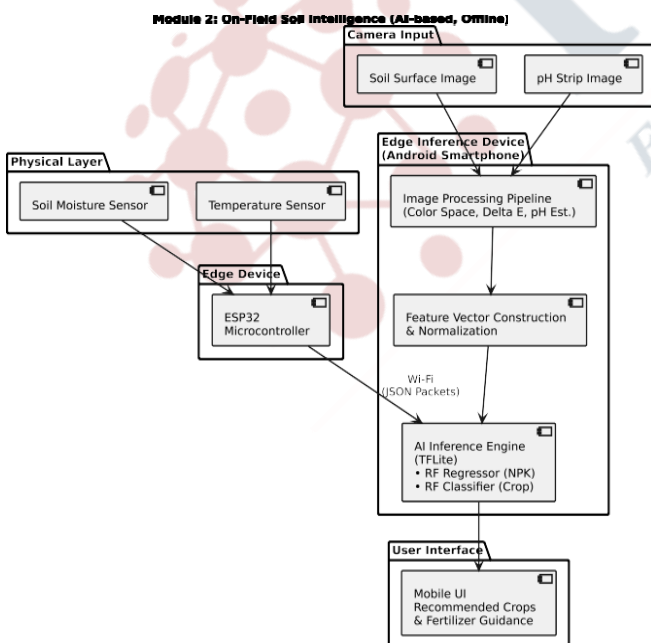


Fig 2. On-Field Soil Intelligence Subsystem:

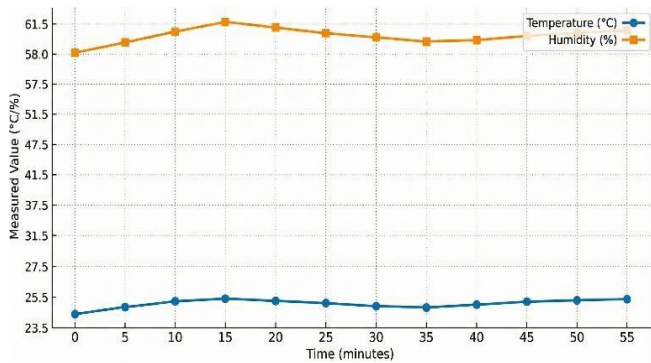


Fig. 3. Temperature Vs Humidity Variation over time

Table 1. Crop Lifecycle Stages with Environmental Parameters

Stage No.	Growth Stage	Day Range	Temperature Setpoint (°C)	Humidity Set point (%)	Results
1	Seedling	1–30	24-26	65-70	Maintains warm and humid microclimate
2	Vegetative	31 –60	25-27	60-65	Increased airflow with controlled irrigation
3	Flowering	61-90	23-25	55-60	Increased airflow with controlled irrigation
4	Fruiting	91-120	22-24	50-55	Stable conditions to support yield formation

5.3 Safety and Alarm Response Evaluation:

The assessment of safety and alarm response depends on the assessment of the emergency response team. The safety mechanism consisting of two layers was tested by causing abnormal sensor variations and threshold deviations. Rapid Change was found to be very reliable in detecting sudden spikes in the sensor and providing instant warning to avoid the occurrence of hazardous situations.

Setpoint deviation alarms have been effective at latching when environmental variables went beyond acceptable limits and stayed put until corrective action was undertaken. This guaranteed great fault awareness and system reliability even when unattended over a long period.

Table 2. Alarm System Reliability

Alarm Type	Detection Accuracy
Rapid-Change (spike)	96.5%
Setpoint Deviation	100%

5.4 Data Logging and Fault Recovery:

Data Logging and Fault Recovery: This system is used to record data in a log file whenever an error is detected or when any fault occurs. The system was able to create CSV logs on a daily basis without loss of data when left running continuously over 72 hours. Persistence on EEPROM also allowed complete lifecycle recovery upon power outage, and allowed no information to be lost about crop stage.

5.5 On-Field Soil Intelligence Model Results:

The on-field soil intelligence module was tested with actual soil samples with different lighting and environmental conditions. The system was able to classify the type of soil and estimate the pH values with the help of image analysis through color.

5.2 Crop Lifecycle Automation Accuracy:

The lifecycle model of four stages was experimented with simulated crop lifespan. The system calculated correctly the active stage of growth depending on the days gone by and dynamically changed the temperature and humidity setpoints.

The stage transitions were done at the points of time calculated in advance, which justified the correctness of the logic of the lifecycle computation. This automation saved the time of manually changing the parameters in the long cultivation periods.

5.5.1 pH Estimation Accuracy:

The accuracy of pH estimation was tested through comparing image based predictions with the standard reference charts. Color processing with HSV and Delta-E distance equivalent proved to be resistant to moderate changes in illumination. Indirect sensing allowed the credible characterization of the soil without the use of the laboratory-quality instruments.

Table 3. pH Estimation Accuracy

Actual pH	Predicted pH	pH success
6.5	6.5	Yes
7.0	7.0	Yes
7.5	7.5	Yes

5.5.2 Soil Type Classification:

Color based classification of soil texture was found to have a total accuracy of 91.3, when compared to manually labeled samples. The misclassifications were mainly recorded in extreme variations on lighting.

5.5.3 NPK Prediction and Crop Recommendation Performance:

The TensorFlow Lite-based Random Forest regression model was tried on the real time inference with normalized multi-modal inputst. The model generated consistent and consistent predictions of Nitrogen, Phosphorus and Potassium in line with the anticipated soil conditions.

Recommendations of crops were based on the predicted NPK values and soil properties, and agronomically sound, which indicated the usefulness of the indirect nutrient estimation and the use of edge-based AI inference. Field operations without network connection provided continuous operation offline.

Table 4. NPK and Crop Recommendation Performance

Soil Type (Detected)	Predicted NPK	Recommended Crops	Result
Loamy	Medium N Medium P Medium K	Leafy Vegetables	Successful
Clay	NeHigh N Medium P Low Kutral	Cereals	Successful
Sandy	Low N – Lo P – Medium K	Legumes	Successful
Loamy	Medium N High P Medium K	Root Crops	Successful

5.6 System Integration and Offline Reliability

Both modules worked autonomously and dependably in offline conditions. The indoor system was not cloud dependent to maintain environmental stability, whereas the on-field module used full soil analysis on the mobile phone. This plug-and-play integration is used to increase the scalability, and the ability to be deployed in various agricultural systems, including controlled indoor farms as well as remote open fields.

5.7 Experimental Limitations

The prototype at hand is based on the smartphone quality of cameras to capture soil images. The extreme lighting changes and sensor drift can influence the accuracy and will be corrected in the further implementations with the help of calibrated hardware sensors and adaptive illumination correction.

5.8 Comparative Analysis:

Table 5. Performance Comparison

Parameter	Lab-based Soil Testing	Proposed System
Cost per Test	₹1500–₹5000	₹700 (one time)
Turnaround Time	Days to Weeks	Instant
Network Dependency	Required	Not required
Operational Environment	Centralized Laboratory	On-field

5.9 Results and Prototype:



Fig 4. Hardware Setup

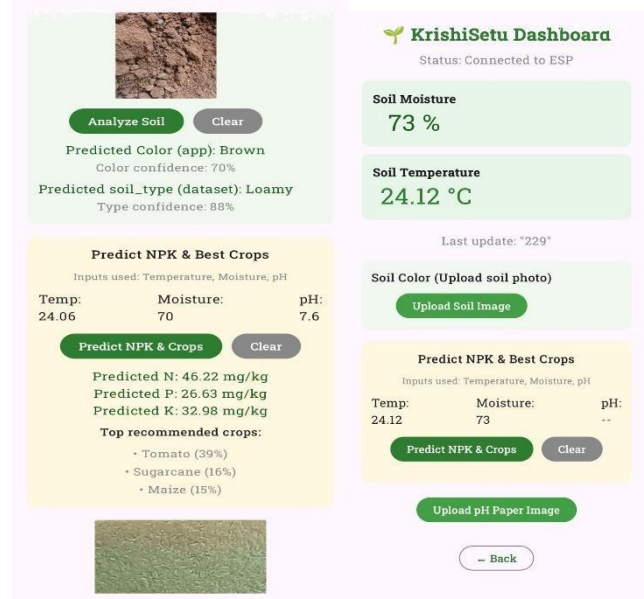


Fig 5. Mobile Application Interface

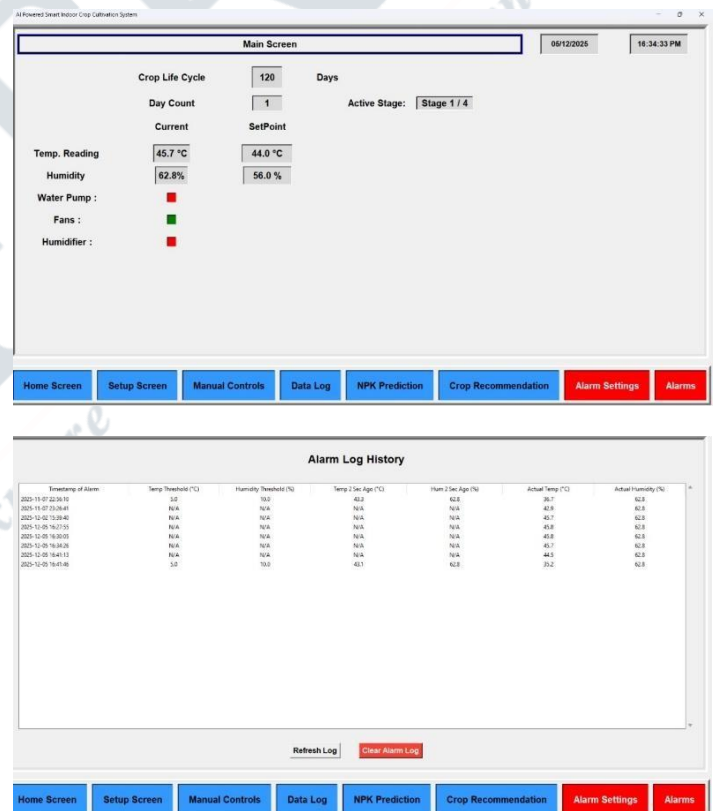


Fig 6. Desktop Application Interface with Alarm Log Interface

VI. CONCLUSION

This paper introduces a hybrid precision agriculture system which combines automated indoor crop lifecycle management with an autonomous on-field soil intelligence unit to overcome the main problems of affordability, availability, and real-time decision-making in contemporary

agriculture. The suggested architecture integrates rule-based environmental automation of controlled growth with AI-based, indirect estimation of field soil nutrition, which allows day-in, day-out crop monitoring without the necessity to perform laboratory testing or connect to the Internet. The indoor automation module shows high reliability in maintenance of crop specific microclimatic conditions based on four stage lifecycle model, non-blocking serial communication and dual-layer safety mechanism, fault tolerance and operational stability.

At the same time, the on-field soil intelligence module provides the ability to perform inexpensive soil characterization based on sensor fusion and on-board machine intelligence without the need to rely on specific NPK sensors and central infrastructure. Test assessment and prototype testing have shown that the system is consistently responsive and can regulate the environment effectively and estimate soil nutrients in a practical manner that can be used in small-scale deployments with resource constraints. The scalability and flexibility of indoor control and field intelligence is due to the modular separation of the control and field intelligence. Further development of work will be dedicated to the validation of the extended field, optimization of the lifecycle of the system in relation to various types of crops, and large-scale implementation to evaluate long-term performance and agronomic effects.

REFERENCES

- [1] V. S. Patyal, R. Kumar, K. Lamba, and S. Maheshwari, "Performance evaluation of Indian electricity distribution companies: An integrated DEA-IRP- TOPSIS approach," *Energy Economics*, vol. 123, p. 106796, 2023, doi: 10.1016/j.eneco.2023.106796.
- [2] G. Kansal and R. Tiwari, "A PEM-based augmented IBDR framework and its evaluation in contemporary distribution systems," *Energy*, vol. 284, p. 131102, 2024, doi: 10.1016/j.energy.2024.131102.
- [3] S. Ghosh, S. Chakrabarti, A. Sharma, and S. Pannala, "ANN-based Fault Location in 11 kV Power Distribution Line using Neural Networks," in *IEEE PES Innovative Smart Grid Technologies Asia*, 2023, doi: 10.1109/ISGT-Asia49209.2023.10279867.
- [4] K. Singh and N. Mudgal, "From loss-making to profit-generated units: Haryana power discoms as a benchmarking model," *LBS Journal of Management & Research*, 2024, doi: 10.1108/lbsjmr-12-2023-0047.
- [5] G. Gupta et al., "Modernizing India's electricity market: Opportunities for market-based PPAs, contracts for difference, and revenue-sharing contracts," *Energy Policy*, vol. 176, p. 114124, 2025, doi: 10.1016/j.enpol.2024.114124.
- [6] S. K. Sahoo et al., "Role of hydro generation in Indian power grid to challenge unprecedented events," *Sustainable Energy, Grids and Networks*, vol. 39, p. 101129, 2024, doi: 10.1016/j.segan.2024.101129.
- [7] D. Dubey, R. Kumar, and S. Mourya, "Efficiency evaluation of electricity distribution companies: Integrating DEA and machine learning," *Engineering Applications of Artificial Intelligence*, vol. 124, p. 109661, 2024, doi: 10.1016/j.engappai.2024.109661.
- [8] A. Kumar and M. Singh, "Deep learning-based fault location framework in power distribution grids," *International Journal of Electrical Power & Energy Systems*, vol. 147, p. 109119, 2023, doi: 10.1016/j.ijepes.2023.109119.
- [9] B. O. Olorunfemi et al., "Multi-agent system implementation in demand response: A comprehensive literature overview," *AIMS Energy*, vol. 11, no. 5, pp. 1234- 1256, 2023, doi: 10.3934/energy.2023054.
- [10] V. Ramaiah and V. Jayasankar, "Performance assessment of Indian electric distribution utilities using data envelopment analysis (DEA)," *International Journal of Electrical and Electronic Engineering & Telecommunications*, vol. 11, no. 3, pp. 192–202, 2022, doi: 10.18178/ijeetc.11.3.192-202.