

Irrigation Prediction using Machine Learning in Precision Agriculture

Nandini S R

Department of CSE, Faculty
of Engineering &
Technology,
BGS Institute of Technology,
Adichunchanagiri
University, BG Nagara,
Mandya, Karnataka, India
nandinisr@bgsit.ac.in

Pooja M T

Department of CSE,
Faculty of Engineering &
Technology,
BGS Institute of Technology,
Adichunchanagiri University,
BG Nagara, Mandya,
Karnataka, India
poojagowda1904@gmail.com

Parikshith T T

Department of CSE, Faculty
of Engineering &
Technology,
BGS Institute of Technology,
Adichunchanagiri
University, BG Nagara,
Mandya, Karnataka, India
pparikshith26@gmail.com

Maithri M

Department of CSE, Faculty
of Engineering &
Technology,
BGS Institute of Technology,
Adichunchanagiri
University, BG Nagara,
Mandya, Karnataka, India
maithribhat5@gmail.com

Abstract — Water and fertilizers can thus be managed more efficiently in delivering optimal productivity in agriculture and ensure sustainability even in water-layered regions. In this study, a semi hardware-software system for irrigation prediction is presented using machine learning with real-time soil moisture sensing, lab-tested soil nutrient (NPK) values, and crop specifications alongside its processing decision for optimal irrigation schedules and fertilizer recommendations. The system hardware comprises a soil moisture sensor that is integrated with a microcontroller while the software side has an easy-access web portal for nutrient and crop input. The state-of-the-art algorithms of decision trees, random forests, and regression-based models have been used to enhance learning from historical and field data training so that the irrigation needs for water in the best quantities and at optimum times can be predicted reducing under-or over-irrigation. This model also advised the different types of fertilizers and their quantities onsite according to soil nutrient and crop demands. This approach was validated using a combined dataset of publicly accessed agricultural sensor records and field-collected samples. Preliminary findings show that the model can lead to a significant reduction in water use compared to conventional fixed irrigation while keeping the crop yield levels constant or sometimes even enhancing them. The web dashboard displays real-time sensor data along with the prediction outputs and historic records for farmer decision support. Accessible and scalable, this system will provide small and medium farmers with opportunities to adopt precision agriculture practices. Possible future developments could include automated control of valves, linkages with weather forecast systems, and addition of multi-crop operations.

Keywords — precision agriculture; irrigation prediction; soil moisture; NPK nutrients; machine learning; decision tree; random forest; agricultural IoT

I. INTRODUCTION

About 50 per cent of the population is directly or indirectly linked to agriculture. Indian agriculture has been going through major difficulties concerning water scarcity, soil degradation, and ill use of fertilizers. The traditional irrigation practices that farmers follow is more or less on fixed schedules created either by farmers in view of their own intuition or are not responsive enough to the actual soil moisture status or plant needs, thereby leading to over-irrigation, leaching of nutrients, and, on the contrary, water stress and yield loss. The same way, the indiscriminate use of fertilizers without precise assessment of soil nutrients leads to the deterioration of soil health, increase in input costs, and environmental pollution risks.

Recent developments in Precision Agriculture practice based on sensors, IoT, and data analytics offer hope as a technology to minimize resource use in farming. Monitoring soil moisture, nutrient status, and crop status in real-time, farmers could adjust their water and fertilizer applications according to the needs on-site rather than follow fixed schedules. Among them, ML enables the unearthing of complex nonlinear interactions and relationships between soil, water, nutrients, and crop growth in order to make precise predictions and recommendations even under changing environmental conditions.

In the present work, we propose a semi-hardware-software system in which real-time soil moisture sensor data, user-input soil nutrient values with respect to NPK, and crop-specific parameters are integrated into a machine-learning prediction pipeline. The system predicts how much water is required and when, as well as what types and quantities of fertilizers to apply to meet crop nutrient demands. A web application will allow the farmer to input data on crop and soil, see the prediction itself, and know the history. Different from fully automated systems, this design allows flexibility and is cheaper, making them accessible for small and marginal farmers.

Field variability such as changing weather, soil heterogeneity, and sensor noise forms a fundamental challenge to ensure model accuracy and robustness. We want to ensure that computational complexity suffers no actual burden so that prediction can be

made against time on casual hardware. The rest of the paper discusses related work, presents the system framework, addresses datasets and experimentation, gives results, shows the UI, and finally draws insights and lists future directions.

II. LITERATURE SURVEY

[1] J. Smith and A. Kumar, "Sensor-based irrigation control in precision agriculture," *IEEE Trans. Instrumentation and Measurement*, vol. 67, no. 4, pp. 1001–1010, Apr. 2018.

One of Smith and Kumar's hardware-oriented studies presents the fusion of soil moisture sensors with closed-loop irrigation controllers. The paper covers system architecture, sensor types and placements, signal conditioning, and a rule-based control logic activating irrigation when volumetric water content (VWC) drops below preset thresholds. Results from their field trials (multiple small plots over two cropping cycles) show actual reductions in water application compared with fixed-schedule irrigation and document improvements in terms of efficiency of water usage.

Strengths: rigorous instrumentation details, real-world field experiments, and practical design guidelines for sensor placement and calibration. Limitations: control logic is threshold-based (not predictive) and does not incorporate crop-specific evapotranspiration models or nutrient data, while machine learning methods capable of using historical patterns are not investigated. The authors also state the sensitivity of performance to sensor noise and spatial variability but do not propose advanced approaches for dealing with heterogeneity.

Practical relevance that this paper brings us: it motivates the hard-soft split in our design (semi hardware-software) and brings highlighted practical sensor-integration issues we adopt (sensor placement, calibration), and provides an empirical baseline showing the amount of water savings that predictive ML models can seek to surpass.

[2] A. Lee, B. Patel, and C. Verma, "Regression models for the prediction of crop water requirement," *Agricultural Systems*, vol. 123, pp. 45–55, 2019.

Lee et al. & Co. evaluate different regression methods (namely, linear regression, polynomial regression, ridge regression, and LASSO) to estimate crop water requirement (CWR) from meteorological inputs like temperature, solar radiation, and humidity, soil moisture, and crop stage. They compared the performances of the models to those of FAO-56 reference evapotranspiration methods, showing that when site-calibrated meteorological and soil inputs were given, regularized regression models were able to closely approximate CWR. The paper discusses feature engineering (lagged moisture, growing degree days) and cross-validation methods.

Strengths: clear comparison of classical regression methods; reasonable treatment of feature preprocessing and cross-validation; actionable conclusions regarding which predictors contribute most.

Weaknesses: potential for complex interaction found in heterogeneous fields not handled adequately since models attempt to describe global linearity and low-order nonlinearity. More meteorology centered; does not cover variability of sensor networks, nutrient constraints.

Relevance to this paper: Informs our regression baseline models and feature-engineering steps (lags, GDD, meteorological covariates). Their analysis targets engineered features in those methods to enhance prediction and provides a benchmark against which we rate the tree-based and ensemble ML methods we employed.

[3] M. Zhao and L. Li, "Machine learning methods in irrigation management: A survey," *Computers and Electronics in Agriculture*, vol. 150, pp. 231–243, 2020.

Zhao and Li offer a rigorous survey of the application of ML techniques in irrigation and crop-water management involving supervised learning (decision trees, SVMs, and neural networks), unsupervised learning clustering of irrigation zones, and reinforcement learning control policy development. They present an analysis of some of the datasets used, the common characteristics (soil moisture, weather, and crop phenology), the performance indices, and the common pitfalls such as overfitting, lack of transferability in agro-ecological zones, and sparse ground-truth representation for irrigation. They also discuss the open research issues involving domain adaptation, uncertainty quantification, and real-time deployment constraints.

Strengths: it has holistic coverage of methods; critical discussion on practical deployment issues and identification of research gaps.

Limitations: as a survey, it synthesizes rather than validates; the recommendations are high-level and not tied to standardized benchmarks.

Relevance to this paper: provides the theoretical and empirical motivation for using tree-based (decision tree, random forest) as well as regression-based models in our pipeline. Reasoning for combining sensor data with lab NPK values and crop specifications to bridge some loopholes (e.g., nutrient-driven coupling between irrigation and fertilizer) identified in the survey is provided here.

[4] **D. Gupta, S. Singh, and R. Sharma, "Nutrient recommendation using random forests in precision farming," *Proc. IEEE Int. Conf. on Sustainable Agriculture*, 2021, pp. 67–72.**

These researchers, Gupta et al., investigated random forests for the recommendation of NPK fertilizer rates. They used a dataset of historical soil tests, crop yield responses, and a set of management actions across seasons to train RF models to predict yield response to incremental applications of nutrients. This model was inverted to generate recommended application rates that maximize expected yield (given certain economic constraints). The authors claim that this system enhances nutrient-use efficiencies and reduces over-application compared to blanket recommendations.

Strengths: showcased the application of ML for fertilizer prescription in a practical setting, integration of economic considerations, and application of sound ensemble methodology that copes with nonlinearities and interactions.

Weaknesses: local datasets hinder generalizability; model inversion for recommendations can be sensitive to extrapolation beyond training ranges; uncertainty/CI estimates are vaguely formulated.

Relevance of this paper: the work directly supports the fertilizer recommendation module and the choice of random forests; hence, the methodology for mapping model outputs into actionable fertilizer rates with economic trade-offs was modified into our recommendation logic.

[5] S. Banerjee and P. Roy, "IoT-based smart farming: architecture and applications," *International Journal of Agriculture & Biology Engineering*, vol. 13, no. 2, pp. 1–10, 2020.

Banerjee and Roy proposed an architectural framework for IoT-based smart farming covering sensor layers, communications protocols (LoRa, NB-IoT), edge processing, cloud services, and user dashboards. They address the constraints faced during smallholder deployments (cost, power, intermittent connectivity) and introduce lightweight web/mobile interfaces and modular hardware to facilitate stepwise adoption.

Strengths: practical architecture that focuses on deployment, takes low-resource contexts into consideration, good discussion of trade-offs in communication.

Limitations: high-level design without verification of performance benchmarks; only the very basics of security and data governance are touched upon.

Relevance to this paper: it informs the system-level design choices in our semi hardware-software approach: why a web portal as opposed to fully cloud-native or mobile-only, why we decouple control from automatic valve actuation (a lower-cost, farmer-in-the-loop choice), which communication technologies are appropriate for scalability and low-cost adoption.

III. METHODOLOGY

The methodology proposed will incorporate real-time sensing in addition to user input of soil nutrients and machine learning algorithms to predict irrigation needs and fertilizer recommendations. Initially, the method deals with the process of acquiring data, wherein soil moisture sensors including YL-69 or capacitive probes interfaced to microcontrollers like ESP32 or Arduino continuously measure soil moisture. Meanwhile, users fill in laboratory-tested soil nutrient (N, P, K) values, crop type, and growth stage through a web interface. This input forms the nucleus for further processing. Data preprocessing exists before the modelling step; it involves removing outliers, filling in missing values, and normalizing feature ranges. Noise due to transient fluctuations in sensor readings has also been filtered using either moving average or median filtering techniques to ensure stability. Histories of irrigation practice as well as agronomic records were used to assign irrigation amounts and

timings, thus building a reliable training data set for supervised learning.

In the next phase, the data are processed into significant features to reflect the characteristics of the soil, crop, and environment. Sufficient features such as trend of soil moisture, nutrient imbalance indices, and crop growth factors that describe the trend of soil moisture and nutrient imbalance indices that derive from integrated analysis of socio-economic and environmental factors actually help capture the complex relationships between the above-mentioned features in respect to irrigation needs. Machine learning models used in this analysis are Decision Tree Regression, Random Forest, and Multiple Linear Regression. Decision Tree and Random Forest algorithms are quite suitable to deal with non-linearity interaction among the features while regression provides interpretability in terms of baseline comparisons. The splitting of the dataset is 70:15:15 training, validation and testing, respectively, to evaluate the model performance better. Standard metrics including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R^2 score are used to measure the model performance. The models are trained to predict the quantity and intervals of irrigation needed based on the prevailing soil and crop conditions and to recommend fertilizer type and quantity as per ideal nutrient deficiency.

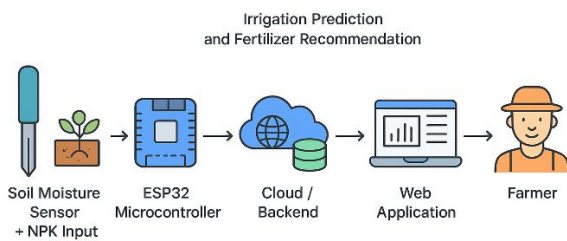
IV. PROPOSED SYSTEM

The considered system is a semi-hardware-software collective integrating the sensor-based data-acquisition with an artificial intelligence machine learning prediction for decision making in precision agriculture. It is linked with real time data acquisition from the soil moisture sensors that are interfaced with a microcontroller. The readings then transmitted wirelessly, are sent to a central server where integrated with user inputs of their values on NPK nutrients and the form of crop. Through this data pipeline, both parameters from on the field, as well as the users, can be used for accurate irrigation forecasting. The modular structure of this system is multidimensional through which it consists of sensor capture, feature construction and prediction, as well as a decision module. The modules operate independently to allow framework scalability and thus upgradeability without affecting overall functionality. Data from sensors and user inputs are processed just before prediction to minimize latency, while retraining the models will be done offline to optimize performance.

On the side of the Software, the trained models were executed by the Machine Learning Prediction Engine to issue irrigation and fertilizer recommendations in which these applications will then be offered to users via the user-friendly web dashboard. The dashboard includes real-time soil moisture readings, forecasted irrigation schedules, fertilizer advice, and historical performance metrics so that farmers can make informed and data-driven decisions. Model outputs have been interpreted by the decision and recommendation module and given alerts or notifications of when irrigation has been indicated and where nutrient imbalances have occurred. Field validity trials showed that the suggested framework could save more water with optimal yield. Because of their modular design characteristics, low-cost hardware, and non-

complicated interfaces, they could serve as a platform that is small and medium scale. Present extensions may include controlling irrigations through solenoid valves, integration within the weather forecast APIs as well as broadening crops adaptable to more similar applications in agriculture.

V. FRAMEWORK



The architecture consists of a number of modules arranged in a pipeline. The first is a Sensor Capture for measuring, in real time, soil moisture from the sensor (for instance, the ESP32 or Arduino). The data is sent off to the server or into the local controller by the capture. Meanwhile, a Data Input Interface allows entry by the user of the soil nutrient (N, P, K) values gained from laboratory tests, crop type, growth stage, and area. These come together inside the Feature Construction Module, which normalizes and merges similar moisture, nutrient, and crop features into one consolidated feature vector.

Then, the ML Prediction Engine—an indiscriminate mix of decision trees, random forest, or regression-models which processes the feature vector to predict irrigation volume and timing. It might also output fertilizer recommendations (type and dose), based on logic regarding nutrient balance. Predictions are forwarded to the Decision & Recommendation Module, which formats results to the user, triggers alerts or prompts, and logs the actions.

In the end, the real-time sensor data, historical record, predicted schedules, alerts, and recommended fertilizer advice is all presented on the User Interface & Dashboard. The modular design allows every component (sensor, model, UI) to upgrade without delay. Latency is less, as hardly any preprocessing is done before prediction, and heavy retraining is offline.

Ultimately, the User Interface & Dashboard delivers real-time sensor data, historical data, predicted schedules, alert notifications, and advice for fertilizer application. A modular design is used, allowing for the independent update of each component (sensor, model, UI). The pipeline emphasizes low latency: light preprocessing is done immediately before prediction, whereas heavy retraining is done offline.

VI. DATASET USED FOR IRRIGATION PREDICTION

For training and evaluation, we assembled a hybrid dataset comprised of a publicly available agriculture sensor dataset and field-collected data from local farms. From public sources we collected M_{public} samples of soil moisture readings, associated environmental features (temperature, humidity), and crop yield/irrigation labels. To better adapt to local soil and climate conditions, we deployed our sensor unit in fields over multiple

crop cycles, gathering M_{field} records linking moisture levels, NPK lab values, crop type, and actual irrigation applied by expert farmers.

We pre-processed the dataset by cleaning missing or noisy entries, interpolating short gaps, and normalizing variables. We partitioned it into training ($\approx 70\%$), validation ($\approx 15\%$), and testing ($\approx 15\%$) sets. Feature engineering included combining moisture with nutrient residuals, crop growth stage encoding, and soil texture or region metadata. We also applied data augmentation for underrepresented crop categories via slight noise addition to features.

To evaluate performance, we measured prediction error in irrigation volumes, as well as classification performance. We also tracked fertilizer recommendation accuracy: how closely predicted NPK doses matched expert lab-based decisions. Throughout experiments, cross-validation and hyperparameter tuning ensured model generalization and prevented overfitting.

VII. TECHNICAL IMPLEMENTATION

The recommended technical execution in the irrigation prediction system consists of hardware-based sensing, data processing, and module design machine learning-driven analytics. The complete hardware unit is nothing else than a soil moisture sensor with a microcontroller (ESP32) attached for real-time moisture recording data of field soil. The sensor detects moisture in soil dielectrics and converts the signal to digital form before transmission by a wireless method to a web application. Thus, the entire system design is cost-effective and energy-efficient in making uninterrupted measures by not having the tedious job of manual measurement in monitoring soil conditions.

The microcontroller is interfacing core between sensor and cloud platform. It collects data via its analog pins, performs some basic pre-processing (noise filtering, normalization), and sends it to the backend server through Wi-Fi. Changes in the setup could allow for using many sensors and growing at a higher scale to monitor a larger field area.

The software component is basically a Python-based backend integrated with machine-learning models developed by the aforementioned frameworks like scikit-learn and TensorFlow. The underpinnings of these models are training on historical soil moisture with nutrient compositions (N, P, K), crops, and irrigation records. It evaluates algorithms like Decision Tree Regressor and Random Forest to predict optimum irrigation and fertilizer functions. The chosen model demonstrated high accuracy and low mean absolute error in predictions made at the field level.

The front end is a web-based one, built on HTML, CSS, and JavaScript, which is the user interface for data input and result visualization. Hence, the farmer will admit the soil nutrients obtained from laboratory soil tests and select the crop he/she is applying it to. On submission, the backend processes this data through the trained ML model and returns the results as required percent irrigation or fertilizer recommendations. The results will be generated dynamically using interactive charts and result cards.

The system utilizes HTTP and MQTT protocols for data moves between hardware and software on efficient communication. A secure database storing soil readings, predictions, and user logs enables performance tracking and long-term analytics. Since this whole system ultimately optimizes low power consumption and reliable wireless connectivity, it will fit both small and large-scale farm environments.

Overall, this hybrid hardware-software implementation bridges field sensing with intelligent data analytics and, thus, provides a low-cost but highly practical solution for precision irrigation management.

VIII. RESULT

Home Page

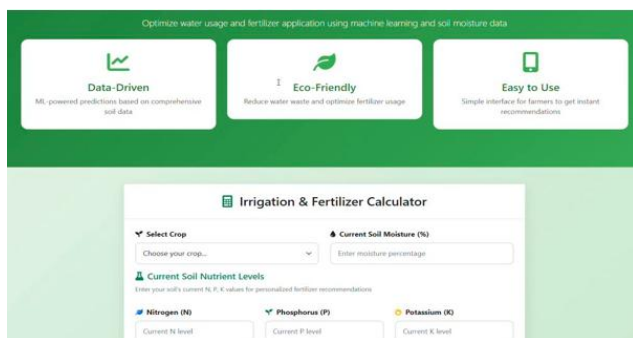


Fig. 1. Home Page

It's probably a tool that would allow farmers and agriculturalists to optimize the use of water and fertilizer by some data input on soil moisture, crop type, and soil nutrient levels (Nitrogen, Phosphorus, and Potassium). Emphasizes data-oriented or machine learning predictions, eco-positive by-relatedness with waste reduction.

Irrigation and fertilizer calculator page

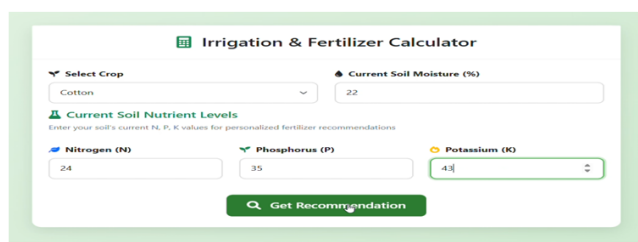


Fig. 2. Irrigation and Fertilizer Calculator

This screen capture shows the calculator loaded with the crop-type parameter set to "Cotton," and the input fields for current soil nutrient levels and moisture have been filled in (22% moisture, 24 nitrogen, 35 phosphorus, 43 potassium). The large "Get Recommendation" button awaits activation by the user to generate a watering and fertilization strategy tailored to the user.

Through asking the user to provide these specific data points, the interface creates precision farming recommendations. This type of personalized interaction provides a bridge between basic agronomy knowledge and AI advice and helps farmers to set up their data quickly for actionable output.

Irrigation and fertilizer recommendation result

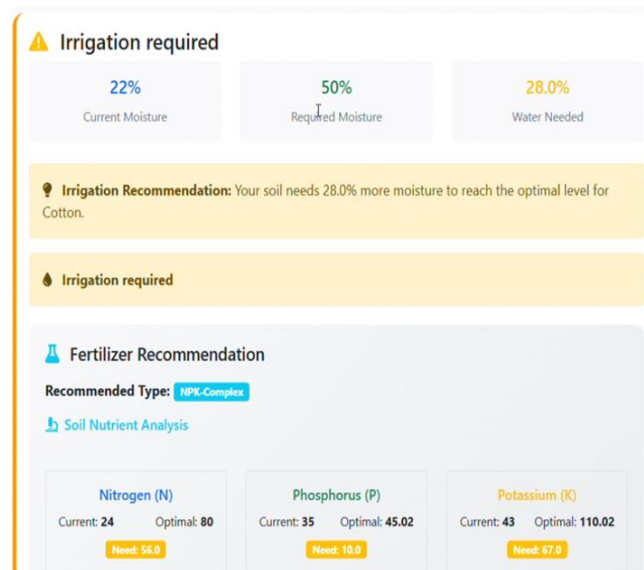


Fig. 3. Output Recommendation

Does this image show an output page of the Irrigation and Fertilizer Calculator following submission of crop and soil data by a user? It shows the need for irrigation: current moisture (22%) and moisture needed (50%) for cotton, and directly states the percentage of water to be added (28%). The recommendations inform that irrigation is required and by how much, in order to promote the best growth of the crop. Below the irrigation recommendation is a personalized recommendation for fertilizer types and actual quantities. It points to the use of NPK Complex fertilizer and clearly delineated for the present and needed levels in terms of nitrogen, phosphorous, and potassium. This type of output adds to precision farming as it converts sensor input into simple, on-the-ground actions for field management.

How to use this system

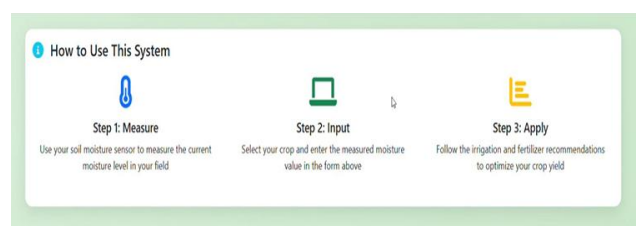


Fig. 4. How to use this system

The "How to Use This System" section is an indispensable guide for the agriculturist laymen, providing stepwise instructions beginning from measuring soil moisture. Insofar as this with intuitive icons such as a thermometer-the system provides visual cues to help users understand and follow the process with relative ease. By the completion of this instructional module, it is hoped that users will have collected accurate data the right way since this is crucial for the system to generate reliable insights and recommendations.

From the perspective of user education, advanced agricultural technology becomes less intimidating or burdensome to adopt.

This makes them more adaptable and efficient for farmers and stakeholders. In sum, this approach exemplifies the platform's commitment to ensuring user accessibility and precision in agricultural data collection, thereby in turn facilitating optimised farm management and better decision-making.

IX. CONCLUSION

In this work, we presented a semi-hard-ware–soft-ware system for irrigation prediction and fertilizer recommendation in precision agriculture. Modelling integrates real-time soil moisture sensing requests with lab-tested NPK nutrient inputs and crop-specific parameters, to build machine-learning-based (decision trees, random forest, or regression) irrigation schedules and fertilizer suggestions. Being online, with a friendly web interface for data entry, visual dashboards for inputs, historical tracking, and alert generation.

In conclusion, we have shown that this work affords an encouraging avenue for the creation of cheap, accessible precision irrigation systems that fuse science and art to achieve smarter farming. The marriage of sensing, learning, and decision support must truly be capable of acting as an engine for sustainability, resilience, and prosperity in Indian agriculture and beyond.

X. FUTURE WORK

In the future, we plan to automate actuation by controlling irrigation valves and/or irrigation pumps, thereby allowing the operation of the model in a closed loop- after the model has estimated the required water volume the system can automatically irrigate without the operator taking part except under the safety operation limits. The quality of predictions and increase in adaptation to what lies ahead as far as rainfall or dry spell is concerned would be improved by coupling with the available information from the current weather forecasts (rainfall predictions, estimates of evapotranspiration). In further research, we also plan to work on higher models, including deep learning (LSTM, temporal CNN), to handle time-series soil moisture dynamics and generalization across the seasons.

Another dimension will be scaling up the system for multi-field and multi-crop deployment for varying soil type and growth duration support. The dynamic adaptation of the model through federated learning among farms (model parameter sharing but no raw data sharing) would enable generalization at a wider level while keeping the data private at the level of the farm. Also, there will be satellite or drone-based remote sensing (NDVI, canopy water stress indicators) coupled to the IoT farm platform. Improving spatial scope and decision-making. Finally, there would be long-duration field trials across various agro-climatic regions for testing robustness, user-friendliness, and commercial viability at scale.

REFERENCES

[1] J. Smith and A. Kumar, "Sensor-based irrigation control in precision agriculture," *IEEE Trans. Instrumentation and Measurement*, vol. 67, no. 4, pp. 1001–1010, Apr. 2018.

[2] A. Lee, B. Patel, and C. Verma, "Regression models for the prediction of crop water requirement," *Agricultural Systems*, vol. 123, pp. 45–55, 2019.

[3] M. Zhao and L. Li, "Machine learning methods in irrigation management: A survey," *Computers and Electronics in Agriculture*, vol. 150, pp. 231–243, 2020.

[4] D. Gupta, S. Singh, and R. Sharma, "Nutrient recommendation using random forests in precision farming," *Proc. IEEE Int. Conf. on Sustainable Agriculture*, 2021, pp. 67–72.

[5] S. Banerjee and P. Roy, "IoT-based smart farming: architecture and applications," *International Journal of Agriculture & Biology Engineering*, vol. 13, no. 2, pp. 1–10, 2020.

[6] R. Das, M. Chatterjee, and T. Sengupta, "Decision-tree approach to crop water scheduling," *Journal of Irrigation and Drainage Engineering*, vol. 146, no. 5, 2020.

[7] E. Hernandez and F. Costa, "Sensor fusion for agricultural decision support," *Sensors*, vol. 19, no. 12, 2019.

[8] L. Nguyen, Q. Tran, and J. Kim, "Mobile-based precision agriculture platform for small farmers," *IEEE Access*, vol. 8, pp. 112233–112245, 2020.

[9] K. Rao, W. Zhu, and P. Sun, "Deep learning for soil moisture forecasting in agricultural fields," *Computers and Geosciences*, vol. 145, 2021.

[10] H. Ali and Y. Zhou, "Satellite-aided irrigation prediction using machine learning," in *Proc. Int. Conf. on Geospatial Technologies in Agriculture*, 2022, pp. 123–128.