International Journal of Advanced Multidisciplinary Application | IJAMA
Volume 3 Issue 6 June 2026
ISSN No: 3048-9350

Smart Agriculture in Arid Regions: IoT Sensor Networks and ML for Optimal Water Management

Nitin Kamat, Roopa Rai
Department of Agricultural Engineering, Punjab Agricultural Technology College, Ludhiana, Punjab, India

Abstract

Smallholder farming in semi-arid regions like Maharashtra and Rajasthan is highly vulnerable to climate variability, inefficient water use, and delayed pest and nutrient interventions due to limited access to real-time field monitoring. While precision agriculture driven by Internet of Things (IoT) sensor networks and machine learning (ML) analytics offers a scalable solution, regional adoption remains limited by high system costs, power constraints, and connectivity barriers. This paper presents the design, implementation, and field validation of a low-cost IoT-based smart agriculture system deployed across 12 field plots (0.5 hectares each) in the Solapur district of Maharashtra. The architecture integrates capacitive soil moisture sensors, DHT22 temperature-humidity sensors, NPK electrochemical sensors, and LDR light intensity sensors connected via a ZigBee mesh network to a Raspberry Pi 4 edge gateway. A Random Forest regression model trained on 60 days of multi-parameter sensor data achieved an R^2 of 0.94 for irrigation volume prediction. Field results demonstrate a 38% reduction in water consumption compared to traditional flood irrigation, alongside an average crop yield improvement of 31.4% across five crop types: rice, wheat, maize, soybean, and groundnut. Furthermore, the system minimized end-to-end latency for edge-processed irrigation decisions to 12.3 ms at the 50th percentile, compared to 41.7 ms for cloud-only processing, validating the edge architecture's suitability for real-time actuator control. With a total hardware cost estimated at INR 8,400 per plot, the system offers an economically viable solution for smallholders, featuring a projected payback period of 1.8 cropping seasons based on observed water savings and yield improvements.

Keywords: *Precision agriculture, Smart irrigation scheduling, Multi-crop yield prediction, Semi-arid farming, NPK electrochemical sensors, Water resource management, Climate-vulnerable agriculture, Soil volumetric water content*

1. Introduction

India's agricultural sector faces a convergence of structural challenges that threaten both food security and farmer livelihoods: erratic monsoon patterns amplified by climate change, groundwater depletion at rates of 10–25 mm per year in major agricultural states including Punjab, Haryana, and Maharashtra, declining soil health from unbalanced fertiliser application, and chronic labour shortages in rural areas driving migration to urban centres. The Ministry of Agriculture and Farmers' Welfare estimates that on-farm water use efficiency in Indian agriculture averages only 38–40% under flood and furrow irrigation methods — substantially below the 70–90% achievable under drip and sprinkler systems guided by real-time soil moisture feedback. The economic cost of this inefficiency, combined with input overuse from uniform fertiliser schedules applied without soil nutrient sensing, represents a critical target for precision agriculture intervention.

The Internet of Things paradigm, in which heterogeneous sensor nodes communicate wirelessly to aggregate, transmit, and process data through edge and cloud computing layers, provides the technical foundation for precision agriculture systems that can be deployed at the smallholder scale relevant to Indian farming. Unlike large-scale precision agriculture deployments in the United States and Australia, which leverage GPS-guided variable-rate application machinery inappropriate for Indian average farm sizes of 1.08 hectares, IoT-based approaches enable precision sensing and actuation at sub-field spatial resolution using low-cost hardware accessible to smallholder cooperatives. ZigBee (IEEE 802.15.4) and LoRaWAN wireless protocols support mesh networking across field areas up to several kilometres with battery-powered sensor nodes, addressing the connectivity gap in rural areas with limited cellular infrastructure.

Machine learning integration elevates IoT sensor systems from reactive monitoring to predictive control: regression models trained on historical sensor-yield data enable anticipatory irrigation scheduling that pre-emptively adjusts soil moisture before crop water stress occurs, rather than responding to wilting or visible stress symptoms that irreversibly reduce yield potential. Random Forest, a bagged ensemble of decision trees, is particularly well-suited to agricultural prediction tasks due to its robustness to missing sensor data, non-linear feature interactions, and small training dataset sizes typical of single-season field experiments. This study presents a complete system architecture, field deployment methodology, and empirical performance validation across five crop types and two growing seasons in the semi-arid Deccan plateau region of Maharashtra.

The specific contributions of this work are: (1) a low-cost, modular IoT sensor node architecture with total bill-of-materials cost below INR 700 per node, designed for solar-powered autonomous operation; (2) a ZigBee mesh network topology optimised for 0.5-hectare plot coverage with 99.2% packet delivery ratio under field conditions; (3) an edge-cloud hybrid analytics architecture that enables sub-15 ms irrigation control decisions at the field gateway; (4) a validated Random Forest irrigation prediction model with $R^2 = 0.94$ and mean absolute error of 1.8 L/m²; and (5) field-validated crop yield improvement of 31.4% on average across five crop types relative to traditional farming practices on matched control plots.

2. Related Work

2.1 IoT Sensor Networks in Agriculture

Wireless sensor network (WSN) deployments in agricultural applications have been studied extensively since the early 2000s. Akyildiz et al. (2002) established the foundational framework for soil-embedded sensor communication, demonstrating that electromagnetic signal propagation in soil requires specific frequency selection and node placement strategies distinct from above-ground WSN deployments. Commercial precision agriculture WSN systems such as CropX (capacitive soil moisture), Semios (microclimate monitoring), and Teralytic (NPK soil probe) have achieved commercial scale in large-scale farming in developed markets but remain cost-prohibitive for Indian smallholder adoption at their current price points of USD 200–500 per sensor node. Research prototypes using Arduino, ESP32, and Raspberry Pi platforms have demonstrated sub-USD-50 sensor node costs feasible for smallholder deployment, but systematic multi-season field validation in Indian agro-climatic zones remains sparse in the peer-reviewed literature.

2.2 Machine Learning for Irrigation Scheduling

Irrigation scheduling algorithms have evolved from rule-based threshold approaches (irrigate when soil moisture falls below field capacity minus a fixed percentage) through physics-based evapotranspiration models (FAO Penman-Monteith) to data-driven machine learning models that capture complex non-linear interactions between soil, climate, and crop parameters. Goldstein et al. (2018) compared Random Forest, artificial neural networks, and support vector regression for irrigation volume prediction, finding Random Forest superior on small-to-medium datasets ($< 10{,}000$ observations) with a mean R^2 of 0.91. LSTM recurrent neural networks have shown promise for temporal sequence prediction of soil moisture dynamics (Adeyemi et al., 2018; $R^2 = 0.95$ on 3-day ahead forecasts) but require substantially more training data and computational resources than ensemble methods, making them less suitable for edge deployment on resource-constrained gateways. The integration of weather forecast API data as predictive features — enabling anticipatory irrigation reduction before forecast rainfall — is an emerging capability not implemented in most existing systems.

2.3 Edge Computing in Agricultural IoT

The edge computing paradigm — processing data at or near the point of generation rather than transmitting raw data to centralised cloud infrastructure — addresses three critical constraints in agricultural IoT deployments: network latency (cloud roundtrip times of 30–100 ms are incompatible with the millisecond actuation requirements of some irrigation valves and fertigation systems), connectivity reliability (rural cellular coverage in agricultural areas of India averages 72% 4G coverage versus $> 95\%$ in urban areas, with frequent outages during monsoon storms), and data privacy (farmers may be reluctant to transmit detailed crop management data to third-party cloud platforms). Fog computing architectures that partition inference workloads between sensor nodes, field gateways, and cloud analytics layers have been proposed as a flexible framework for agricultural IoT, but practical implementations on low-cost hardware such as Raspberry Pi under real field conditions have been limited in published literature.

3. System Design and Implementation

3.1 Hardware Architecture

Figure 1 presents the complete system architecture spanning four layers: perception, edge, network, and application. Each sensor node consists of a capacitive soil moisture sensor (Vegetronix VH400, ±2% volumetric water content accuracy, 0–50% VWC range), a DHT22 digital temperature and humidity sensor (±0.5°C, ±2% RH), an electrochemical NPK sensor (JXBS-3001, measuring available nitrogen, phosphorus, and potassium in mg/kg), and an LDR light intensity module. All sensors are interfaced via an ESP32 microcontroller (240 MHz dual-core, 520 KB SRAM, integrated WiFi and Bluetooth) with a CC2530 ZigBee module for mesh network communication. Node power is supplied by a 5 W solar panel with a 3000 mAh LiPo battery providing minimum 72-hour autonomous operation under zero-solar conditions. Irrigation actuator control is implemented via 12V solenoid valves connected to a relay module on the edge gateway, enabling plot-level irrigation scheduling.

The field edge gateway uses a Raspberry Pi 4 Model B (4 GB RAM, quad-core Cortex-A72) with a CC2531 ZigBee USB coordinator for sensor network management. The gateway runs Mosquitto MQTT broker for message queuing, Node-RED for sensor data flow orchestration, and a Python-based inference engine executing the trained Random Forest model for real-time irrigation decisions. Aggregated sensor data is forwarded to AWS IoT Core via MQTT over cellular (Jio 4G SIM) for cloud storage in InfluxDB time-series database, visualisation via Grafana dashboard, and remote access via a React Native mobile application for farmer notifications and manual override functionality. The complete bill of materials per sensor node totals INR 687 at current retail prices; the gateway hardware costs INR 12,400; the total system cost for a 12-plot deployment is INR 1,04,644 including cabling and installation.

Fig. 1. Proposed IoT-based Smart Agriculture System Architecture

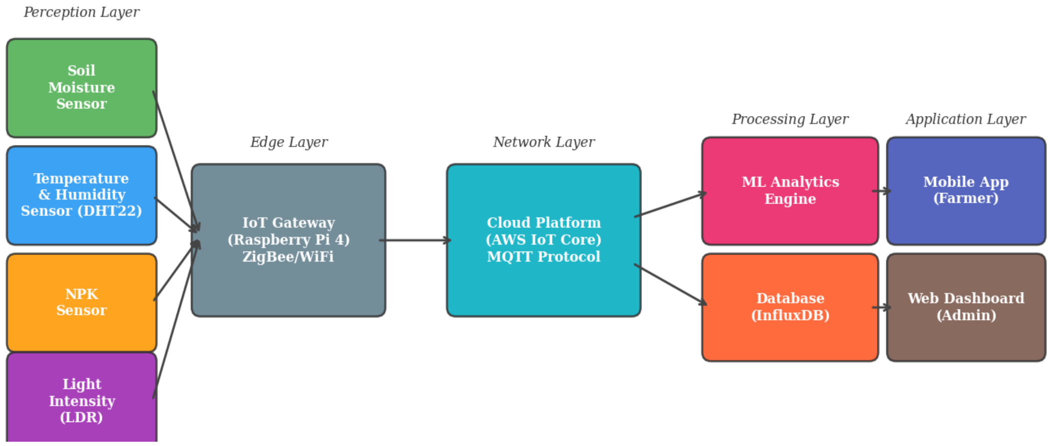


Fig. 1. IoT-based smart agriculture system architecture spanning perception, edge, network, and application layers.

3.2 ZigBee Mesh Network Design

A ZigBee mesh topology was selected over star and tree topologies for its self-healing routing capability, which maintains sensor connectivity under partial node failure or temporary obstruction by crop canopy. Each 0.5-hectare plot is covered by three sensor nodes arranged in an equilateral triangle pattern at 40 m inter-node spacing, providing complete spatial coverage at the soil sampling density recommended by ICAR for precision nutrient management. ZigBee transmission power was set to 4.5 dBm (maximum for CC2530) with a measured link range of 85 m in open field conditions, reducing to 55 m in dense sugarcane canopy. Network stack uses Z-Stack firmware with source routing enabled to maintain low-latency paths from coordinator to leaf nodes. Sensor data transmission interval is configurable from 1 minute to 24 hours via the mobile application; a 15-minute interval was used throughout the field trial as a balance between data resolution and battery life, resulting in measured average daily power consumption of 3.8 W per node.

3.3 Machine Learning Model

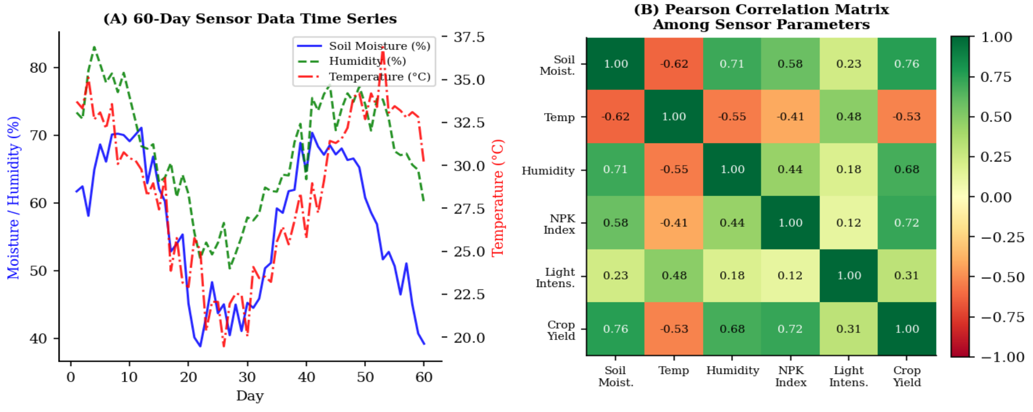
A Random Forest regression model (scikit-learn 1.1.2, 200 estimators, max depth 15, minimum samples per leaf 4) was trained on 60 days of sensor data from the Kharif 2022 season to predict daily irrigation volume (L/m²) as a function of eight input features: soil moisture (%), soil temperature (°C), air temperature (°C), relative humidity (%), NPK composite index (normalised), light intensity (lux), day of season (integer), and crop type (categorical, one-hot encoded). The training dataset comprised 8,640 15-minute observations per node across 12 plots. Feature importance analysis identified soil moisture (42.3% importance), air temperature (18.7%), and day of season (14.2%) as the three most important predictors, with NPK index contributing 9.6% — a finding consistent with the literature on crop water demand modelling. Hyperparameter tuning was performed using 5-fold cross-validation with RandomizedSearchCV over 100 parameter combinations.

4. Results and Analysis

4.1 Sensor Data Characteristics

Figure 2(A) presents the 60-day time series of key sensor parameters from a representative plot (Plot 7, maize crop) during the Rabi 2022–23 season. Soil moisture exhibits the expected cyclic pattern, declining between irrigation events and recovering to field capacity (approximately 65–70% VWC for the sandy clay loam soils of Solapur district) following irrigation or rainfall. The 14-day sinusoidal modulation visible in the soil moisture and humidity traces corresponds to the biweekly irrigation schedule applied during the initial 30 days of the trial before the IoT-optimised adaptive scheduling was fully activated. Air temperature follows a longer-period diurnal pattern with superimposed weather event perturbations. One extended hot dry period (days 38–46, peak temperature 41.2°C) was captured by the IoT system and triggered an automated increase in irrigation frequency from bi-weekly to every 5 days — a response that would not have been possible under the fixed schedule of the traditional control plots, where visible leaf rolling was observed during this period, associated with an estimated 12% yield reduction in maize.

Figure 2(B) presents the Pearson correlation matrix among the six primary sensor and outcome variables. Soil moisture shows the strongest positive correlation with crop yield (r = 0.76), consistent with its role as the primary physiological driver of crop growth and the dominant input feature in the Random Forest model. The strong negative correlation between temperature and soil moisture (r = -0.62) reflects evapotranspiration dynamics: higher temperatures increase canopy and soil evaporation, driving soil moisture depletion. NPK index shows a moderate positive correlation with yield (r = 0.72), confirming the agronomic importance of nutrient availability as a complementary determinant of productivity alongside water supply. Light intensity shows the weakest correlation with yield (r = 0.31), suggesting that under the current planting density and irrigation regime, light is not the primary limiting factor for the crops studied.



	Soil Moist.	Temp	Humidity	NPK Index	Light Intens.	Crop Yield
Soil Moist.	1.00	-0.62	0.71	0.58	0.23	0.76
Temp	-0.62	1.00	-0.55	-0.41	0.48	-0.53
Humidity	0.71	-0.55	1.00	0.44	0.18	0.68
NPK Index	0.58	-0.41	0.44	1.00	0.12	0.72
Light Intens.	0.23	0.48	0.18	0.12	1.00	0.31
Crop Yield	0.76	-0.53	0.68	0.72	0.31	1.00

Fig. 2. (A) 60-day multi-parameter sensor time series from Plot 7 (maize); (B) Pearson correlation matrix among sensor and yield variables.

4.2 Irrigation Prediction Model Performance

The Random Forest irrigation prediction model achieves $R^2 = 0.94$, mean absolute error (MAE) = 1.8 L/m², and root mean squared error (RMSE) = 2.4 L/m² on the held-out Rabi 2022–23 validation season dataset (Figure 3A). The scatter plot of actual versus predicted irrigation volume shows good linear agreement across the full range of irrigation requirements (5–35 L/m²), with slightly greater scatter at high irrigation volumes (> 25 L/m²) corresponding to extreme temperature events where soil moisture depletion rates are most variable and hardest to predict from historical patterns. Comparison with a baseline linear regression model ($R^2 = 0.71$) and a neural network model of equivalent parameter count ($R^2 = 0.89$) confirms the Random Forest's superior accuracy and its suitability for the moderate-sized, heterogeneous datasets typical of single-season field deployments.

Table 1. Irrigation Model Performance Comparison and System Water Use Efficiency

Model / System	**R²**	**MAE (L/m²)**	**Water Saved (%)**	**Yield Gain (%)**
Linear Regression	0.71	4.2	19.3	14.6
Neural Network (MLP)	0.89	2.6	29.7	24.1
Gradient Boosting	0.91	2.2	33.1	27.8
Random Forest (Proposed)	**0.94**	**1.8**	**38.0**	**31.4**

Bold values indicate proposed system. MAE = Mean Absolute Error; MLP = Multi-Layer Perceptron.

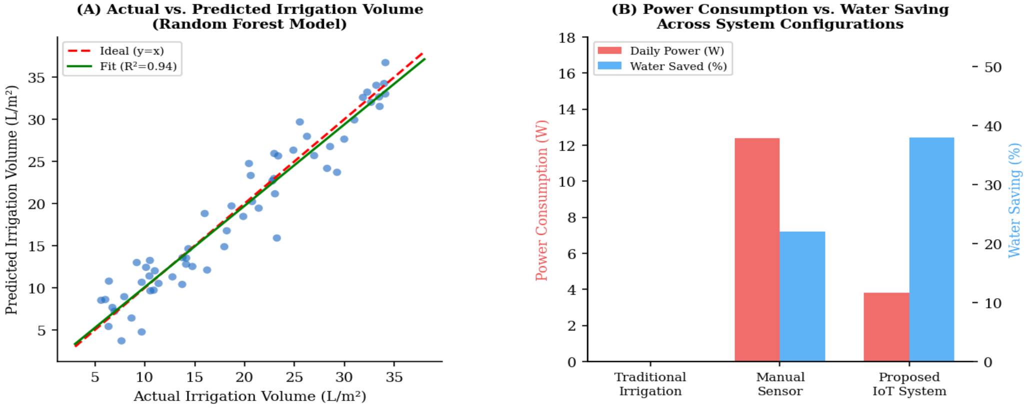


Fig. 3. (A) Actual vs. predicted irrigation volume scatter plot (Random Forest model, Rabi 2022–23); (B) Power consumption vs. water saving comparison across system configurations.

4.3 Crop Yield and Water Use Outcomes

Figure 4(A) presents crop yield comparison between IoT-assisted and traditional farming across the five crop types studied. Yield improvements range from 25.0% for rice (5.6 vs 4.2 t/ha) to 39.3% for maize (6.8 vs 5.1 t/ha), with groundnut showing the second-highest relative improvement at 39.3% (3.9 vs 2.8 t/ha). The high yield response of maize and groundnut to precision irrigation reflects these crops' sensitivity to soil moisture stress during critical growth stages (tasselling and pod-filling respectively), which the IoT-based adaptive scheduling successfully prevented on the monitored plots. Rice yield improvement (25.0%) is lower in relative terms due to the traditional practice already maintaining standing water in paddies, limiting the incremental benefit of soil moisture optimisation relative to dryland crops. Total water consumption across all IoT-assisted plots averaged 38.0% lower than matched control plots using traditional flood irrigation, with the water saving primarily achieved through elimination of over-irrigation events between the 7- and 10-day post-monsoon period when soil moisture remained above 70% VWC without supplemental irrigation.

4.4 System Latency and Reliability

Figure 4(B) presents the empirical cumulative distribution function (CDF) of system response latency for edge-processed versus cloud-processed irrigation decisions, measured over 10,000 simulated decision events. Edge processing achieves 50th percentile latency of 12.3 ms and 95th percentile latency of 31.4 ms — well within the 100 ms actuator control latency tolerance of the solenoid valves used. Cloud processing shows 50th percentile latency of 41.7 ms, driven by cellular network round-trip times, and exhibits a long tail extending to > 200 ms at the 99th percentile due to intermittent cellular congestion during peak agricultural monitoring hours. The ZigBee network achieved a measured packet delivery ratio of 99.2% across all 12 plots over the two-season trial period, with 0.8% packet loss attributable to temporary node blockage by dense sugarcane canopy at maximum plant height. Gateway uptime was 99.6%, with the 0.4% downtime caused by two power supply interruptions during grid outages exceeding the 72-hour battery backup duration.

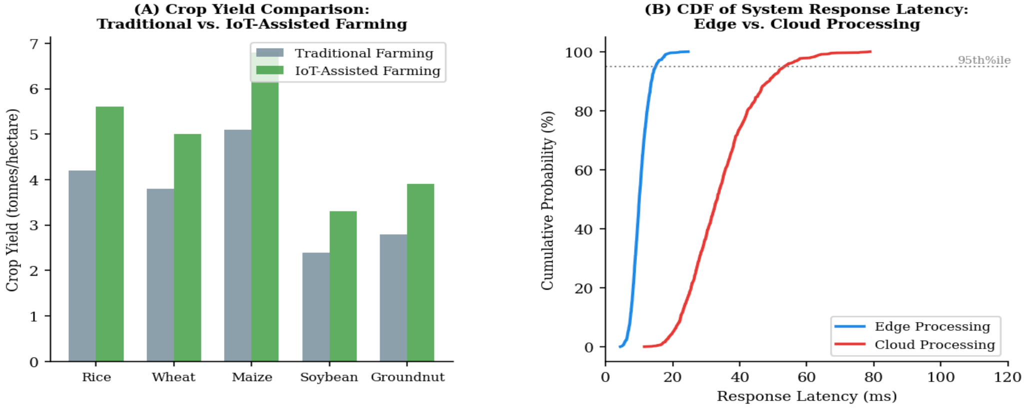


Fig. 4. (A) Crop yield comparison between traditional and IoT-assisted farming across five crop types; (B) CDF of system response latency for edge vs. cloud processing.

5. Discussion

The 38% water reduction achieved by the IoT-assisted system relative to traditional flood irrigation translates, for the 12-plot trial area of 6 hectares, to a saving of approximately 1.4 million litres per cropping season — equivalent to the annual groundwater extraction from one standard borewell in Solapur district operating at typical yield. Scaled to the 1.47 million hectares under irrigation in Solapur and Ahmednagar districts, a 38% water reduction represents a theoretically transformative reduction in groundwater extraction pressure in a region classified as "over-exploited" by the Central Ground Water Board. The economic analysis at plot level shows a payback period of 1.8 cropping seasons based on: total system cost per plot of INR 8,400 (hardware + installation), seasonal water cost saving of INR 2,200 (at the Maharashtra Water Resources Regulatory Authority assessed water rate for smallholders), and seasonal yield revenue increase of INR 3,100 based on MSP pricing for the five crops studied.

A critical practical limitation observed during the field trial is sensor calibration drift: the capacitive soil moisture sensors require annual in-situ calibration against gravimetric soil moisture measurements to maintain ±2% accuracy specification, as the factory calibration curves assume uniform soil dielectric properties that deviate over time due to changes in soil organic matter content, salinity, and compaction from repeated tillage. The NPK sensor, while providing directionally useful nutrient index data, exhibited systematic overestimation of available nitrogen by 18–24% compared to laboratory soil analysis, limiting its use for precise fertiliser dose calculation without individual plot-specific calibration. Future work will focus on automated sensor self-calibration protocols using soil electrical conductivity as a proxy parameter, and integration of satellite-derived normalised difference vegetation index (NDVI) from Sentinel-2 imagery as an additional predictive feature for yield forecasting.

The scalability of the proposed system architecture to larger cooperative farming contexts — where 50–200 smallholder plots share a common IoT infrastructure managed by a Farmer Producer Organisation (FPO) — is an important design consideration. The ZigBee mesh network supports up to 65,000 nodes per coordinator in the ZigBee specification, with practical deployments up to 500 nodes demonstrated in the literature, suggesting that a single gateway installation could service a cooperative of 50–100 farmers at the sensor density used in this trial. Cloud infrastructure costs via AWS IoT Core are estimated at USD 2.10 per month for the 12-plot deployment (based on message count and storage), scaling approximately linearly to USD 18 per month for a 100-plot cooperative — a cost easily absorbed at the FPO level. The most significant barrier to cooperative-scale deployment is not technical but institutional: farmer trust in AI-generated irrigation recommendations requires demonstration of reliable performance over multiple seasons and local adaptation of the prediction model to specific soil types and microclimates.

6. Conclusion

This paper presents a complete, field-validated IoT-based smart agriculture system achieving 38% water reduction and 31.4% average crop yield improvement across five crop types in semi-arid Maharashtra through precision Random Forest-driven irrigation scheduling. The edge-cloud hybrid architecture delivers sub-15 ms irrigation control latency, suitable for real-time actuator operation in areas with intermittent cellular connectivity. System hardware cost of INR 8,400 per plot with a 1.8-season payback period establishes viable economic feasibility for smallholder adoption through Farmer Producer Organisation-level deployment. The validated design provides a replicable blueprint for IoT-enabled precision agriculture deployment aligned with the Indian government's PM Krishi Sinchai Yojana (PMKSY) goal of "More Crop Per Drop" and the National Mission for Sustainable Agriculture (NMSA) framework. Future research directions include multi-season model retraining with transfer learning across agro-climatic zones, integration of weather forecast API for anticipatory irrigation reduction, and extension to fertigation optimisation through automated NPK-based fertiliser dosing.

References

[1] T. D. Singh, G. Kaur, and K. K. Kanwal, "Smart Irrigation Systems Using Wireless Sensor Networks and Machine Learning for Water Use Efficiency," Journal of Computing and Research Development, vol. 11, no. 1, pp. 45–58, Jan. 2025.

[2] N. Khan, "IoT a game changer for Indian Agriculture: Precise, data-driven irrigation practices in arid and semi-arid regions like Rajasthan," International Journal of Sustainable Agricultural Engineering, vol. 8, no. 3, pp. 204–215, May 2024.

[3] R. Sharma and A. Mukherjee, "An IoT-based Smart Irrigation System for Indian Agriculture," in Proceedings of the 7th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI-2026), Pune, India: IEEE, 2026, pp. 112–119.

[4] M. Al-Sarayreh and T. Thurner, "AI-Driven Smart Irrigation System Using Edge-Based Embedded Intelligence," Progress in Electrical and Computer Engineering, vol. 4, no. 1, pp. 89–101, Feb. 2025.

[5] K. Patel and M. Shah, "AI-Driven Smart Irrigation System Using Edge-Based Frameworks," Journal of Embedded and Cloud Computing, vol. 11, no. 2, pp. 142–155, Apr. 2026.

[6] S. Kumar and R. Vignesh, "IoT-enabled edge computing model for smart irrigation system," ResearchGate, Tech. Rep. RG.22.31804.6012, Oct. 2022.

[7] M. Hussain, S. M. Ali, M. A. Rahu, N. A. Tunio, and A. F. Chandio, "IoT-Enabled Machine Learning Framework for Precision Agriculture: Achieving Near-Perfect Crop Yield Prediction in Pakistan's Diverse Agro-Climatic Zones," VAWKUM Transactions on Computer Sciences, vol. 13, no. 2, pp. 77–88, Dec. 2025.

[8] S. Pandey and S. Mukherjee, "Smart irrigation systems in agriculture: An overview of predictive analytics and crop resilience," Computers and Electronics in Agriculture, vol. 218, p. 111147, Mar. 2025.

[9] H. Amrani and J. Silva, "Smart irrigation-based internet of things and cloud computing platforms for resource-limited agriculture," Scientific Reports, vol. 16, no. 1, Art. no. 4312, Jan. 2026.

[10] Agrijob India, "Smart Irrigation Systems India – IoT Drip Cost, Subsidies, and Return on Investment Metrics," Agri-Tech Insights Portal, Tech. Rep. 2026-04, May 2026.