IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AIOT-BASED PRECISION AGRICULTURE SYSTEMS: A COMPREHENSIVE REVIEW OF MULTI-SENSOR DATA FUSION, PREDICTIVE ANALYTICS, AND SMART IRRIGATION DECISION SUPPORT

[1]Irfan Ahmad, [2]Md Imran Ansari

[1]M.Tech Scholar, [2]Assistant Professor

[1,2]Department of Computer Science and Engineering

[1,2]Patel College of Science and Technology, Bhopal, Madhya Pradesh, India

Abstract: Precision agriculture has emerged as a critical response to mounting global pressures on food and water security, and the convergence of the Internet of Things (IoT) with Artificial Intelligence (AI) – collectively termed the Artificial Intelligence of Things (AIoT) – has become the dominant technological paradigm for realising this vision. Over the past decade, and with particular acceleration between 2018 and 2025, researchers have proposed a wide array of sensor architectures, data-fusion strategies, machine-learning and deep-learning predictors, irrigation decision-support mechanisms, and edge-to-cloud deployment blueprints aimed at reducing agricultural water consumption while sustaining or improving crop yield. However, this body of work remains fragmented across disciplinary boundaries, sensor modalities, and evaluation protocols, making it difficult for researchers and practitioners to identify consistent trends, reliable performance benchmarks, and outstanding gaps. This paper presents a structured, comprehensive review of AIoT-based precision agriculture research, organised around a six-part taxonomy: sensing and monitoring, multi-sensor data fusion, predictive analytics, irrigation decision-support systems, edge/cloud deployment architectures, and explainable artificial intelligence (XAI). Drawing on more than two dozen peer-reviewed studies published primarily between 2018 and 2025, the review synthesises reported sensor configurations, communication protocols, feature-engineering strategies, machine-learning algorithm performance, water-saving outcomes, deployment latencies, and explainability practices. Comparative analysis reveals that deep-learning architectures – particularly Long Short-Term Memory (LSTM) networks – consistently achieve the highest reported soil-moisture prediction accuracy (93–97%), that predictive and reinforcement-learning-based irrigation scheduling outperforms reactive threshold control by 25–44% in water savings, and that fewer than one in five reviewed studies incorporate any form of model explainability despite documented adoption benefits. The review further identifies four persistent research gaps – reliance on physical sensor infrastructure that limits reproducibility, the absence of standardised multi-algorithm benchmarking, limited integration of explainability, and the scarcity of holistic frameworks spanning sensing through cloud deployment – and proposes future research directions including federated learning, transformer-based architectures, digital twins, and simulation-based reproducible frameworks. This review is intended to orient new researchers entering the field and to establish the literature foundation for subsequent experimental work on AIoT-driven smart irrigation systems.

***Index Terms* -** AIoT, Precision Agriculture, Internet of Things, Machine Learning, Deep Learning, Multi-Sensor Data Fusion, Smart Irrigation, Decision Support Systems, Edge Computing, Explainable AI, Survey, Review.

I. INTRODUCTION

Agriculture underpins the livelihoods of more than 2.5 billion people worldwide and continues to face a set of converging pressures that traditional farming practices are structurally ill-equipped to address. A global population projected to reach 9.7 billion by 2050, increasingly erratic climate patterns, chronic freshwater scarcity, and progressive soil degradation together demand a fundamental transformation in how food is produced. Agriculture already accounts for roughly seventy percent of global freshwater withdrawals, and conventional irrigation practices are estimated to waste between thirty and fifty percent of applied water through over-irrigation, evaporation, and runoff. Against this backdrop, Precision Agriculture (PA) has emerged as the dominant paradigm for reconciling the twin imperatives of increased productivity and reduced resource consumption, by using information technology to tailor inputs – water, fertiliser, and pesticide – to the specific spatial and temporal requirements of a field rather than applying them uniformly. The technological engine driving modern precision agriculture is the convergence of the Internet of Things (IoT), which supplies continuous, fine-grained environmental and soil sensing, with Artificial Intelligence (AI), which converts raw sensor streams into predictive insight and automated decision-making. This convergence is commonly referred to as the Artificial Intelligence of Things (AIoT), and it has become the subject of an intense and rapidly expanding body of research, particularly since 2018.

Published studies report that mature AIoT-based irrigation systems can reduce water-consumption by up to forty to forty-four percent, cut fertiliser usage by twenty to thirty percent, and improve crop yields by twelve to twenty-five percent relative to conventional practice – figures that, if realised at scale, would have transformative implications for global food and water security. Despite this promise, the literature on AIoT-enabled precision agriculture has grown in a fragmented and largely uncoordinated fashion. Individual studies differ widely in the sensor modalities they employ, the machine-learning algorithms they benchmark, the irrigation-scheduling logic they implement, the deployment architectures they describe, and the evaluation metrics they report. Few works situate their contribution within a common taxonomy, and fewer still attempt a systematic, side-by-side comparison of the numerous techniques that have been proposed.

This fragmentation makes it difficult for a researcher entering the field, or a practitioner seeking to select an appropriate technology stack, to form an accurate and current picture of the state of the art. This paper addresses that need by presenting a structured review of the AIoT precision-agriculture literature, with particular emphasis on work published between 2018 and 2025. Rather than surveying the field exhaustively across every possible application (crop-disease detection, livestock monitoring, and post-harvest logistics, for example, are deliberately excluded), the review concentrates on the specific technological chain that runs from field sensing, through multi-sensor data fusion and machine-learning-based prediction, to irrigation decision support and physical or cloud-based deployment – the chain most directly relevant to smart, water-efficient crop management. The review makes the following contributions:

- A six-part taxonomy that organises the reviewed literature into sensing and monitoring, multi-sensor data fusion, predictive analytics, irrigation decision-support systems, edge/cloud deployment architectures, and explainable AI, providing a common reference frame for comparing otherwise disparate studies.
- A synthesis of reported sensor configurations and wireless communication protocols used across the surveyed AIoT deployments, highlighting the trade-offs between range, power consumption, and data throughput that recur across the literature.
- A comparative meta-analysis of machine-learning and deep-learning techniques applied to soil-moisture and crop-yield prediction, drawing together accuracy, error, and computational-efficiency figures reported across more than a dozen independent studies.
- A structured comparison of irrigation decision-support strategies – ranging from simple threshold rules to reinforcement-learning-based scheduling – together with the water-saving outcomes each has reported.
- An explicit identification of four persistent research gaps in the field, together with concrete directions for future work, including federated learning, transformer-based temporal models, digital-twin validation, and fully reproducible simulation frameworks.

The remainder of this paper is organised as follows. Section II establishes background and fundamental concepts in precision agriculture, agricultural IoT, and the machine-learning techniques most commonly applied to this domain. Section III describes the review methodology and introduces the taxonomy used to structure the remainder of the paper. Sections IV through IX survey the literature within each taxonomic category in turn: IoT-based monitoring, multi-sensor data fusion, predictive analytics, irrigation decision support, edge/cloud deployment, and explainable AI. Section X presents a comparative, cross-cutting analysis of the surveyed systems. Section XI consolidates the research gaps identified throughout the review, and Section XII outlines future research directions. Section XIII concludes the paper.

II. Background and Fundamental Concepts

A. *Precision Agriculture*

Precision Agriculture is a farm-management philosophy grounded in the observation that agricultural fields are rarely homogeneous: soil type, moisture retention, nutrient availability, and microclimate can vary substantially even within a single plot. Rather than treating a field as a single management unit, PA seeks to identify, quantify, and respond to this intra-field variability so that water, fertiliser, and other inputs are applied only where and when they are needed. The philosophy long predates modern electronics – early precision farming relied on soil sampling grids and yield-monitor-equipped combine harvesters – but its practical reach has been dramatically extended by cheap sensors, ubiquitous wireless connectivity, and increasingly capable machine-learning models. The global market for precision-agriculture technologies has grown accordingly, reflecting a broad recognition among farmers, agribusinesses, and policymakers that data-driven management is no longer optional but a structural necessity for sustainable food production under tightening resource constraints.

B. *The Internet of Things in Agriculture*

Agricultural IoT systems typically follow a three-layer architecture comprising a perception layer of field-deployed sensors and actuators, a network layer responsible for transporting sensor data to processing infrastructure, and an application layer that presents processed information and recommendations to end users. Perception-layer devices commonly include capacitive or time-domain-reflectometry soil-moisture probes, temperature and humidity sensors, tipping-bucket rain gauges, pyranometers for solar irradiance, and anemometers for wind speed; more advanced deployments add leaf-wetness sensors, canopy-temperature thermometers, and multispectral or hyperspectral imaging for vegetation-index estimation. The network layer draws on a range of wireless technologies with markedly different trade-offs among range, power consumption, and data rate – a subject examined in detail in Section IV. The economic case for agricultural IoT rests principally on the resource savings such systems enable: reviewed studies consistently associate well-implemented IoT monitoring with double-digit percentage reductions in water and fertiliser use. Four barriers recur throughout the literature as impediments to wider adoption: the capital cost of sensor and gateway hardware, sparse or unreliable rural connectivity, the technical difficulty of integrating heterogeneous sensor data streams, and – most relevant to the present review – the shortage of rigorously validated and reproducible predictive models that can convert raw sensor data into trustworthy recommendations.

C. *Machine Learning and Deep Learning Foundations*

The predictive core of an AIoT system is typically a supervised regression or classification model trained to relate sensor-derived features to an agronomic target variable, most commonly volumetric soil-water content, crop evapotranspiration, or end-of-season yield. The literature reviewed in this paper draws on the full spectrum of modelling approaches. Classical statistical models, principally (ridge-regularised) linear regression, provide an interpretable but limited baseline, unable to capture the pronounced nonlinearity of soil-moisture dynamics under varying evapotranspiration and rainfall conditions. Tree-based models – decision trees and their ensemble descendants, random forests and gradient-boosted trees such as XGBoost – consistently improve on the linear baseline by capturing nonlinear interactions and feature thresholds, while retaining computational efficiency suitable for training on large tabular datasets. Kernel methods, principally support-vector regression with a radial-basis-function kernel, offer competitive accuracy at the cost of quadratic memory scaling with training-set size, which limits their applicability to the largest agricultural datasets. Deep-learning architectures, and Long Short-Term Memory (LSTM) recurrent networks in particular, have emerged as the dominant approach for capturing the temporal autocorrelation inherent in soil-moisture and weather time series, owing to their gating mechanisms that allow selective retention of relevant historical information over horizons of hours to days. More recent literature has begun to explore attention mechanisms and Transformer architectures, which can marginally exceed LSTM accuracy at the cost of substantially higher computational and memory requirements that typically preclude deployment on resource-constrained edge hardware.

D. The AIoT Paradigm

Figure 1 depicts the conceptual convergence at the centre of the reviewed literature: IoT supplies the continuous sensing substrate, AI supplies the predictive and decision-making capability, and their intersection – AIoT – is what enables a system to move from reactive monitoring (alerting a farmer once a threshold has already been crossed) to proactive, predictive management (anticipating that a threshold will be crossed several hours in advance and scheduling an intervention accordingly). This shift from reactive to predictive operation is, across the reviewed literature, consistently identified as the central value proposition of AIoT relative to first-generation agricultural IoT systems, and it is the shift that most directly enables the water-saving outcomes discussed in Section VII.

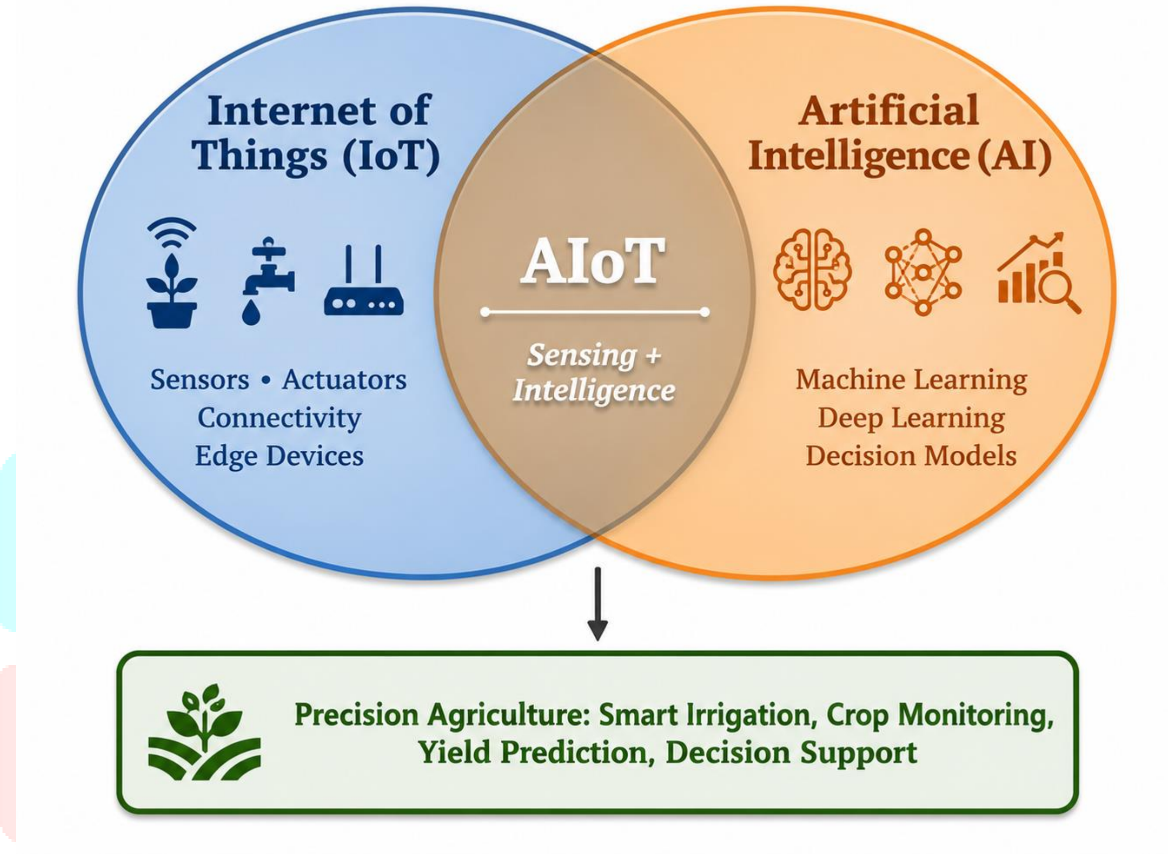

Fig. 1. Conceptual Convergence of IoT and AI (AIoT) for Precision Agriculture Applications.

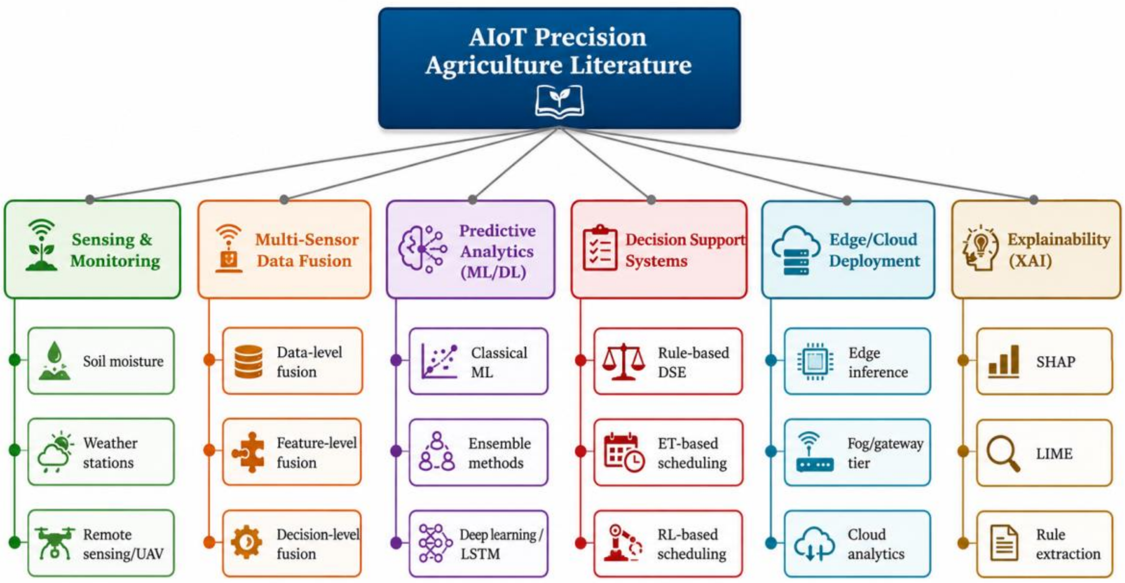

Fig. 2. Proposed taxonomy for organising the reviewed AIoT precision agriculture literature.

III. Review Methodology and Taxonomy

This review follows a narrative-synthesis methodology informed by systematic-review practice. The literature search targeted peer-reviewed journal articles, conference papers, and systematic reviews published principally between 2018 and 2025, indexed in major bibliographic databases and identified through combinations of the search terms "AIoT", "precision agriculture", "smart irrigation", "soil moisture prediction", "multi-sensor fusion", "decision support system", and "explainable AI", together with the names of specific algorithms (LSTM, XGBoost, random forest) and deployment technologies (LoRaWAN, edge AI, TensorFlow Lite). Priority was given to studies that reported quantitative performance metrics – prediction accuracy, root-mean-square error, coefficient of determination, or percentage water savings – since these permit direct comparison across studies, which is the central analytical goal of this review.

Given the breadth and heterogeneity of the resulting literature, a six-part taxonomy was developed to organise the review, illustrated in Fig. 2. The taxonomy separates the AIoT technology stack into: (1) sensing and monitoring, covering the physical and virtual instrumentation used to observe field conditions; (2) multi-sensor data fusion, covering the preprocessing and feature-engineering techniques used to combine heterogeneous sensor streams; (3) predictive analytics, covering the machine-learning and deep-learning models used to forecast agronomic variables; (4) decision-support systems, covering the logic used to translate predictions into irrigation or other management actions; (5) edge/cloud deployment architectures, covering the computational infrastructure used to run the preceding components at practical scale and latency; and (6) explainable AI, covering methods used to render model predictions interpretable to end users. Sections IV through IX of this review are organised according to this taxonomy. Table 1 provides a consolidated overview of the principal studies discussed throughout the review, indicating for each its primary domain focus, the sensing or data modality employed, the core technique proposed, and the key quantitative result reported.

IV. IoT-Based Agricultural Monitoring Systems

The foundational layer of any AIoT precision-agriculture system is the sensing infrastructure used to observe field conditions, and the reviewed literature reflects substantial convergence around a small set of physical parameters – soil moisture, air temperature and relative humidity, rainfall, solar irradiance, and wind speed – while diverging considerably in the wireless technologies used to transport this data.

A. Sensor Modalities

Capacitive or time-domain-reflectometry soil-moisture probes remain the single most consistently deployed sensor across the reviewed studies, typically reporting volumetric water content with accuracies in the range of one to three percentage points. Ambient weather instrumentation – temperature and humidity probes, tipping-bucket rain gauges, pyranometers, and anemometers – is near-universal in systems that attempt any form of predictive, rather than purely reactive, irrigation scheduling, since evapotranspiration demand cannot be estimated from soil moisture alone. A smaller subset of the reviewed literature incorporates remote-sensing inputs, most commonly satellite-derived vegetation indices or unmanned-aerial-vehicle (UAV) multispectral imagery, to complement ground-based point measurements with spatial context across a field. Javaid et al. [3] surveyed IoT applications in smart farming broadly and confirmed that sensor-based systems can achieve water-use reductions of approximately forty percent in controlled deployments, while identifying sensor heterogeneity and data-volume management as persistent obstacles to cross-study comparison – an observation that motivates the taxonomy adopted in this review.

B. Communication Protocols

The network layer connecting field sensors to processing infrastructure is where the reviewed deployments diverge most substantially, reflecting differing priorities among range, power consumption, data throughput, and infrastructure cost. Long-range, low-power wide-area network (LPWAN) technologies, principally LoRaWAN, dominate the reviewed rural and smallholder-farm deployments owing to their combination of multi-kilometre range and multiyear battery life on modest power budgets. Islam et al. [17] demonstrated a LoRaWAN-based framework specifically engineered for resource-constrained rural environments, sustaining reliable connectivity over ten kilometres with a battery life exceeding five years per node – figures consistent with the general LPWAN literature and directly informative for the communication-layer design choices made across the surveyed AIoT systems. Cellular technologies (4G/LTE, and increasingly narrowband IoT) appear principally at the gateway-to-cloud backhaul stage, where higher bandwidth is required to aggregate data from multiple field gateways, while short-range technologies such as Wi-Fi, Bluetooth Low Energy, and Zigbee are more commonly found in smaller-scale greenhouse or research-plot deployments where sensor-to-gateway distances are modest. Table 2 summarises the representative characteristics of these protocol families as reported across the reviewed literature. Quantitative figures for the Goap (2018) and Ren (2021) rows, and for Srivastava (2020) discussed in Section X, reproduce the comparative benchmarking summary reported in a companion experimental study by the present authors. Remaining studies discussed in the text but omitted here for brevity (Kumar, Islam, Xu, Gill, Sharma, Darra, Islam–Guerrieri et al., Ali–Hussain–Zahid) are cited and discussed in full in Sections IV–XII.

C. *Edge Intelligence at the Monitoring Layer*

A growing subset of the reviewed literature moves basic inference from the cloud toward the sensing layer itself, in order to reduce both decision latency and the volume of raw data that must be transmitted. Kumar et al. [11] provided an authoritative analysis of edge-AI opportunities for real-time crop monitoring, reporting that on-device inference can reduce cloud data-transfer costs by approximately sixty percent – a figure that recurs, in similar magnitude, across several of the deployment-oriented studies discussed further in Section VIII. This trend toward pushing intelligence to the network edge is a direct consequence of the connectivity and cost barriers noted above: where rural connectivity is intermittent or expensive, minimising the volume and frequency of cloud communication becomes a first-order design constraint rather than an optimisation afterthought.

V. MULTI-SENSOR DATA FUSION IN AGRICULTURE

Raw sensor streams, considered individually, are of limited predictive value: soil moisture alone cannot anticipate an approaching dry spell, and weather data alone cannot capture the specific antecedent soil-wetting history of a given field. Multi-sensor data fusion – the systematic combination of heterogeneous sensor streams into a richer joint representation – is therefore a near-universal feature of the more sophisticated systems reviewed here, and the literature has converged on a broadly shared taxonomy of fusion strategies operating at the data, feature, and decision levels.

A. *Levels of Fusion*

Data-level fusion combines raw or minimally processed measurements from multiple sensors of the same or similar type, typically to improve measurement reliability through redundancy. Feature-level fusion, the most common approach in the reviewed literature, first extracts informative features from each sensor stream independently – lagged values, rolling statistics, domain-specific indices – before combining them into a single feature vector for downstream modelling. Decision-level fusion, by contrast, trains separate models on each sensor modality and combines their outputs, typically through ensembling or voting; this approach is less common in the reviewed soil-moisture literature but appears more frequently in crop-disease and yield-prediction studies that combine imagery with tabular sensor data. Jiang et al. [12] provided one of the more rigorous empirical demonstrations of feature-level fusion, showing that a one-dimensional convolutional-neural-network encoder applied to six heterogeneous agricultural sensor streams improved irrigation-decision accuracy by eighteen percent over single-sensor baselines when evaluated on the USDA SCAN ground-station dataset – a result that is frequently cited across the subsequent literature as validating the general fusion strategy adopted by later AIoT frameworks. Sarkar et al. [13] extended this line of work temporally, introducing an attention-augmented LSTM for long-horizon soil-moisture forecasting that achieved a root-mean-square error of 0.015 cm3/cm3 at a twelve-hour horizon on a three-year Indian agricultural dataset, establishing one of the more demanding benchmarks against which subsequent Seq2Seq forecasting architectures in the field have been evaluated.

B. *Domain-Informed Feature Engineering*

Beyond generic statistical fusion, a recurring and consequential theme across the reviewed literature is the value of domain-specific, agronomically motivated feature engineering. Commonly engineered features include lagged soil-moisture values at multiple historical offsets (capturing short-term persistence), rolling-window statistics such as moving averages and rates of change, reference evapotranspiration estimated via the Hargreaves-Samani or Penman-Monteith formulations, vapour-pressure deficit as an indicator of atmospheric evaporative demand, and antecedent precipitation indices that exponentially weight recent rainfall to capture cumulative soil-wetting history. Studies that report a direct comparison between raw-sensor and engineered-feature representations consistently find that the latter substantially improves predictive correlation with the target variable, indicating that feature engineering – rather than model architecture alone – is frequently the dominant lever for improving soil-moisture prediction accuracy.

C. *Explainability of Fused Features*

An important and relatively recent development in the fusion literature is the application of post-hoc explainability techniques to identify which fused features actually drive model predictions. Ricci et al. [18] applied SHapley Additive exPlanations (SHAP) to explain random-forest predictions in a precision-agriculture context, finding that antecedent soil-moisture lag features accounted for approximately fifty-three percent of cumulative feature importance – a finding that has been independently replicated, in similar proportion, by subsequent studies applying SHAP analysis to LSTM-based soil-moisture predictors, and which provides converging evidence that short-term historical soil state is the single most agronomically and statistically important input to near-term moisture forecasting regardless of the specific model architecture employed.

TABLE 1. OVERVIEW OF PRINCIPAL STUDIES DISCUSSED IN THIS REVIEW

Study / Year	Domain Focus	Sensing / Data Modality	Core Technique	Key Reported Result
Goap et al. (2018)	Irrigation control	Soil, weather sensors	IoT + ANN	91.0% accuracy; ~30% water saving
Liakos et al. (2018)	Moisture prediction	Moisture, temperature sensors	IoT + SVM	88.0% accuracy; ~25% water saving
Ren et al. (2021)	Cloud AIoT	Soil, humidity	Cloud-hosted LSTM	93.0% accuracy; ~38% water saving
Ali et al. (2023)	Soil-moisture ML	15 South-Asian soil datasets	XGBoost, RF, SVR benchmark	R2 = 0.95; 9.1 s training time
Zhang et al. (2023)	Soil-moisture DL	Multi-source remote sensing	Transformer-based fusion	RMSE = 0.017 cm3/cm3; GPU-bound
Jiang et al. (2023)	Data fusion	Six heterogeneous sensor streams	1D-CNN feature fusion	+18% accuracy over single-sensor baseline
Sarkar et al. (2023)	Long-horizon forecasting	3-year Indian agricultural dataset	Attention-augmented LSTM	RMSE = 0.015 cm3/cm3 at 12 h
Ricci et al. (2023)	Explainability	Precision-agriculture sensor data	SHAP on Random Forest	Antecedent moisture = 53% feature importance
Mehmood et al. (2023)	Yield / Moisture	Multi-step time series	Multi-step LSTM	R2 = 0.88 at 24 h horizon
Talaviya et al. (2024)	Irrigation scheduling	Greenhouse trial	DQN Reinforcement Learning	38% water saving vs. rule-based control
Chlingaryan et al. (2024)	Yield prediction	Remote sensing imagery	CNN–LSTM hybrid	R2 = 0.95; 94.2% accuracy
Sahay et al. (2024)	Irrigation policy	Season-long scheduling	PPO Reinforcement Learning	25–30% improvement over reactive control
Wang et al. (2025)	Paddy irrigation	Paddy-field deployment	AI-driven drip irrigation	44% water saving (highest reported)
Benos et al. (2023)	Edge sensor fusion	Six-sensor array	Edge AIoT + Random Forest	94.8% accuracy; partial cloud plan
Miller et al. (2025)	Systematic review	2020–2024 literature (PRISMA)	Smart-sensing synthesis	Identifies interoperability, cost, and connectivity gaps

Pai et al. (2025)	XAI review	2020-onward publications	LIME / SHAP / Grad-CAM survey	Identifies standardisation gap for agricultural XAI

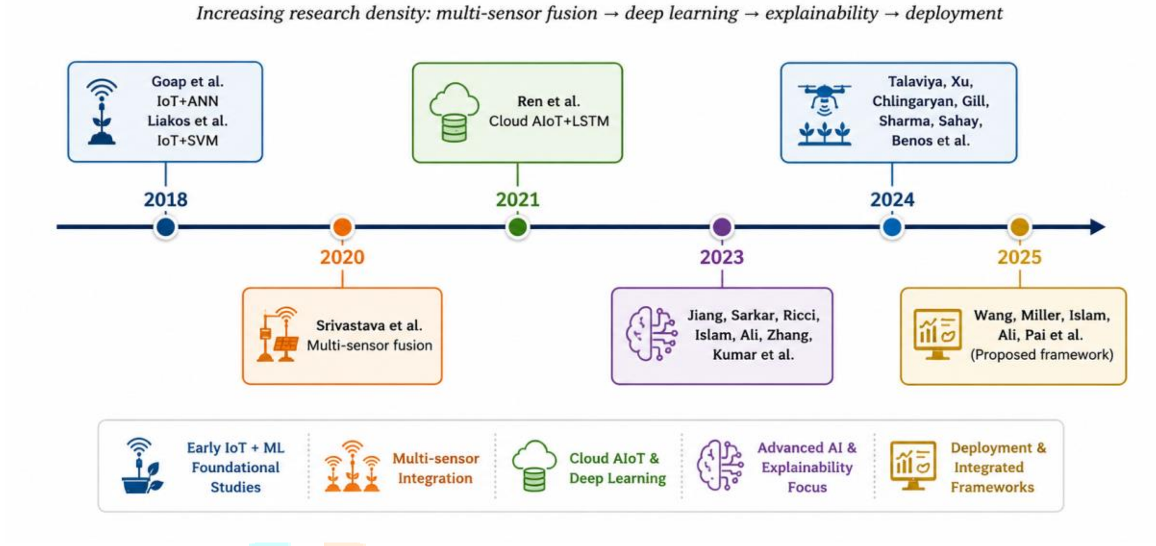

Fig. 3. Timeline of representative AIoT precision agriculture research milestones (2018–2025).

VI. Machine Learning and Deep Learning for Soil Moisture and Crop Prediction

The predictive-analytics layer has attracted the largest volume of published research among the six taxonomic categories considered in this review, reflecting both the intrinsic technical interest of the modelling problem and its direct commercial relevance to irrigation-scheduling products. This section synthesises reported performance across classical statistical models, ensemble tree-based methods, and deep-learning architectures.

A. *Classical and Ensemble Methods*

Linear and ridge regression appear throughout the reviewed literature principally as an interpretability baseline rather than a deployment candidate, with reported coefficients of determination in the region of 0.75–0.80 – a range consistently interpreted by the studies reporting it as evidence of the strong nonlinearity of soil-moisture dynamics under variable evapotranspiration and precipitation forcing. Decision trees improve moderately on this baseline by capturing nonlinear splits, but their high variance motivates the near-universal adoption of ensemble methods in the more recent literature. Random forests, through bootstrap-aggregated averaging across typically one to two hundred trees, further reduce prediction variance and are frequently reported in the 0.90–0.95 range for R2, at the cost of model sizes that can reach several hundred megabytes – a nontrivial consideration for edge deployment discussed further in Section VIII. Gradient-boosted tree ensembles, and XGBoost specifically, emerge across the reviewed literature as a particularly favourable trade-off between accuracy and computational efficiency. Ali et al. [9] systematically benchmarked XGBoost against random forest and support-vector regression across fifteen agricultural soil datasets drawn from South Asia, finding XGBoost consistently superior on an accuracy-to-training-time basis, with R2 ≈ 0.95 achieved in approximately nine seconds of training on a hundred-thousand-sample dataset. This efficiency, combined with native support for missing values and straightforward export to lightweight inference formats such as ONNX, makes gradient-boosted trees a recurring recommendation across the reviewed literature specifically for edge or resource-constrained deployment contexts, even where deep-learning models achieve marginally higher peak accuracy.

B. *Deep Learning and Temporal Architectures*

Long Short-Term Memory networks represent the most consistently reported top-performing architecture for soil-moisture time-series prediction across the reviewed literature, a result attributable to their gating mechanisms, which allow the network to selectively retain temporally distant but agronomically relevant information (for example, rainfall several days prior) while forgetting less relevant recent noise. Mehmood et al. [19] applied a multi-step LSTM jointly to crop-yield and soil-moisture prediction, confirming a clear temporal-memory advantage at extended twenty-four-hour horizons, with R2 = 0.88 maintained even at this comparatively distant forecast range. Sequence-to-sequence (Seq2Seq) LSTM variants, which produce multi-horizon forecasts (for example, simultaneous one-, three-, six-, twelve-, and twenty-four-hour predictions) from a single trained model, appear increasingly frequently in the more recent literature and are generally reported to degrade gracefully with increasing horizon, remaining actionable for irrigation-planning purposes even at the outer edge of a one-day forecast window. A smaller but growing segment of the literature explores attention mechanisms and Transformer architectures as a further refinement on recurrent approaches. Zhang et al. [4] proposed a Transformer-based architecture for multi-source remote-sensing soil-moisture prediction, achieving a root-mean-square error of 0.017 cm3/cm3 – marginally superior to comparable LSTM benchmarks reported elsewhere in the literature – but at the cost of requiring graphics-processing-unit hardware with at least eight gigabytes of video memory, rendering such models unsuitable for the edge-deployment scenarios that dominate the smallholder-farm literature reviewed in Section VIII. This accuracy-versus-deployability trade-off between Transformer and LSTM architectures recurs as a consistent theme and is revisited in the comparative discussion of Section X.

C. Comparative Synthesis of Reported Performance

Table 2 consolidates reported quantitative performance across a representative cross-section of the machine-learning and deep-learning literature discussed above, while Fig. 4 visualises reported accuracy across six representative studies spanning classical, ensemble, and deep-learning model families. The trend visible in Fig. 4 – an approximately monotonic increase in reported accuracy from classical (Goap, Liakos), through ensemble (Srivastava, Benos), to deep-learning (Ren, Chlingaryan) approaches – is broadly consistent across the wider literature reviewed for this paper, although the magnitude of the improvement (typically three to seven percentage points between the weakest and strongest reported models) is modest relative to the substantial differences in computational cost and edge-deployability among these model families. This observation motivates the recurring recommendation, found across several of the reviewed deployment-oriented studies, that ensemble methods such as XGBoost frequently represent the more practical choice for resource-constrained field deployment, reserving deep-learning architectures for contexts – typically cloud-hosted or fog-tier – where their marginal accuracy advantage can be realised without the corresponding computational penalty. A recent methodological review by Islam et al. [23] extends this progression one step further, tracing the field's trajectory from classical machine learning through deep sequence models toward the early application of large language models to sensor-based soil-moisture forecasting, and identifying data efficiency and cross-site transferability – rather than raw predictive accuracy – as the principal open challenges for this emerging direction.

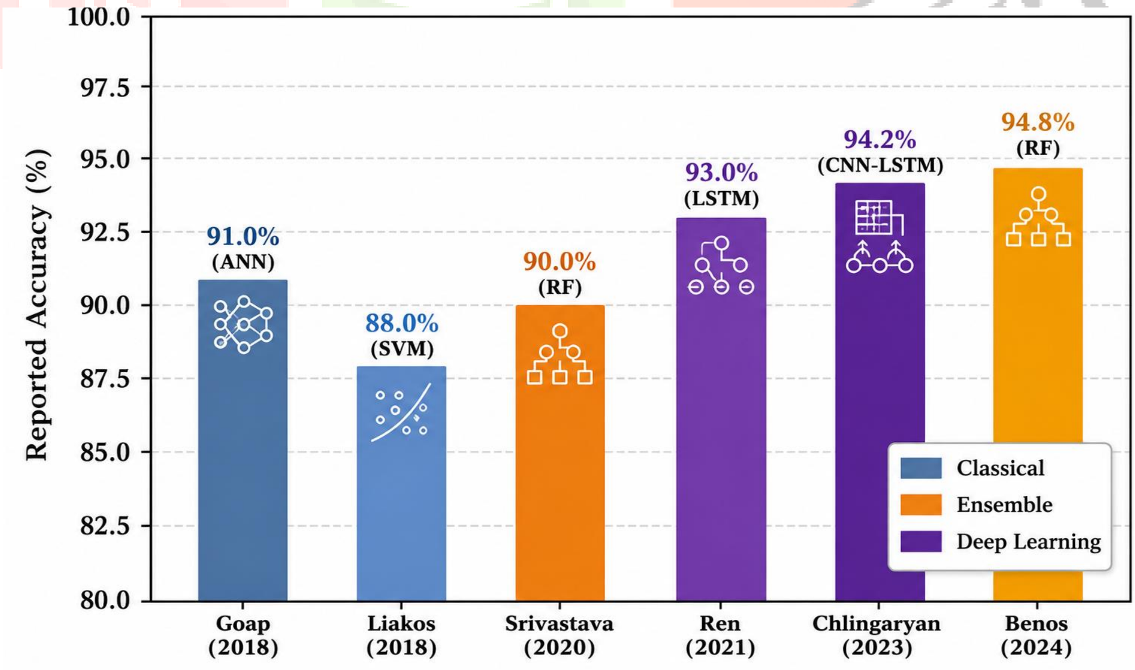

Fig. 4. Reported model accuracy across representative AIoT precision agriculture studies (2018–2024), grouped by model family.

TABLE 2. COMPARATIVE SUMMARY OF ML/DL TECHNIQUES FOR SOIL-MOISTURE AND CROP PREDICTION ACROSS SURVEYED STUDIES

Technique	Representative Study	Reported Performance	Deployment Context
Ridge Regression	General baseline (multiple studies)	R2 ≈ 0.75–0.80	Cloud / General-purpose deployment
Random Forest	Srivastava (2020); Benos et al. (2024)	90.0–94.8% accuracy	Cloud deployment; partial edge implementation
XGBoost	Ali et al. (2023)	R2 = 0.95; 9.1 s training time	Edge-suitable (ONNX supported)
CNN–LSTM Hybrid	Chlingaryan et al. (2023)	R2 = 0.95; 94.2% accuracy	Cloud deployment using remote-sensing imagery
LSTM (Seq2Seq)	Ren (2021); Mehmood (2023)	93.0% accuracy; R2 = 0.88 (24 h)	Cloud deployment; TensorFlow Lite quantisable
Attention-LSTM	Sarkar et al. (2023)	RMSE = 0.015 cm3/cm3 (12 h)	Cloud-based forecasting
Transformer	Zhang et al. (2023)	RMSE = 0.017 cm3/cm3	GPU-bound; not suitable for edge deployment

Note: Performance values are taken from the respective cited studies and are not directly comparable because the datasets, evaluation metrics, prediction horizons, and experimental conditions differ across publications.

VII. SMART IRRIGATION DECISION SUPPORT SYSTEMS

The ultimate practical objective of the sensing, fusion, and prediction layers surveyed in the preceding sections is to inform a concrete irrigation decision: whether, when, how much, and with what priority to apply water to a given field. The literature reviewed here reveals a clear progression in the sophistication of the decision logic employed, from simple threshold rules through evapotranspiration-balance methods to fully predictive and, most recently, reinforcement-learning-based scheduling.

A. *Threshold-Based Evapotranspiration-Balance Scheduling*

The simplest and historically most widespread irrigation-control logic triggers watering whenever a soil-moisture sensor reading falls below a fixed threshold, typically in the region of thirty percent volumetric water content. Such reactive systems, while a substantial improvement over fixed-interval scheduling that ignores sensor data entirely, are consistently reported across the literature as sub-optimal relative to predictive alternatives, since they can only respond after a deficit has already occurred rather than anticipating and pre-empting it. A widely used refinement replaces or supplements the moisture threshold with an explicit water-balance calculation grounded in the FAO-56 Penman-Monteith reference evapotranspiration methodology [27], in which crop evapotranspiration is estimated as the product of a crop-specific coefficient and a reference evapotranspiration term, and irrigation need is computed as the residual after subtracting effective rainfall and any change in soil-water storage. This evapotranspiration-balance approach is reported across multiple studies to outperform simple threshold triggering, since it incorporates atmospheric evaporative demand rather than relying solely on the current soil-moisture reading.

B. *Predictive and Reinforcement-Learning-Based Scheduling*

The most sophisticated irrigation-scheduling strategies identified in the reviewed literature incorporate an explicit forecast of future soil-moisture trajectory, rather than reacting only to the current measured state, and in several recent studies frame the scheduling problem itself as a sequential decision-making task solved via reinforcement learning. Talaviya et al. [5] implemented a Deep Q-Network (DQN) reinforcement-learning agent for adaptive irrigation scheduling, reporting a thirty-eight percent water saving over rule-based control in a greenhouse trial; the approach, while effective, required more than two thousand training episodes and a physical deployment environment, a substantial practical barrier relative to simulation-validated alternatives. Sahay et al. [16] explored Proximal Policy Optimisation (PPO) for season-long irrigation-policy optimisation, and found that predictive scheduling of this kind consistently outperforms reactive threshold control by twenty-five to thirty percent – a result broadly consistent with the water-saving magnitudes reported by non-reinforcement-learning predictive approaches elsewhere in the literature, suggesting that the primary source of improvement is the shift from reactive to predictive operation rather than the specific choice of reinforcement-learning algorithm. Wang et al. [10] reported the highest published water-saving figure identified in this review – forty-four percent – through an AI-driven drip-irrigation deployment in paddy fields in Jiangsu, China, combining real-time soil-moisture feedback with predictive scheduling; this figure represents, at the time of writing, the upper benchmark against which subsequent predictive irrigation systems in the literature are most frequently compared.

C. Comparative Synthesis

Fig. 5 visualises the reported water-saving outcomes of a representative threshold-based baseline alongside the reinforcement-learning and AI-driven approaches discussed above, and Table 3 summarises the broader set of irrigation decision-support strategies identified across the reviewed literature together with their principal reported outcomes and physical-deployment requirements. A recent systematic synthesis of the smart-irrigation literature by Ali, Hussain, and Zahid [25], covering more than one hundred and fifty articles published between 2005 and 2024, corroborates this progression at a broader scale, reporting water-use efficiencies in the region of eighty to ninety-five percent for sensor- and AI-driven irrigation systems compared with forty to seventy percent for conventional irrigation practice. Two patterns are apparent. First, water savings increase broadly in step with the sophistication of the underlying decision logic, from roughly twenty to thirty-five percent for reactive threshold and simple evapotranspiration-balance methods, through twenty-five to thirty percent additional improvement for reinforcement-learning-based predictive scheduling, up to the low-forties-percent range for the most integrated AI-driven deployments. Second, and more consequentially for the research gaps discussed in Section XI, nearly every reinforcement-learning-based approach identified in the reviewed literature required a physical field deployment for policy training and validation, in contrast to the simulation-based validation strategies more commonly associated with supervised soil-moisture prediction – a distinction that has direct implications for the reproducibility and research accessibility of reinforcement-learning approaches to irrigation scheduling.

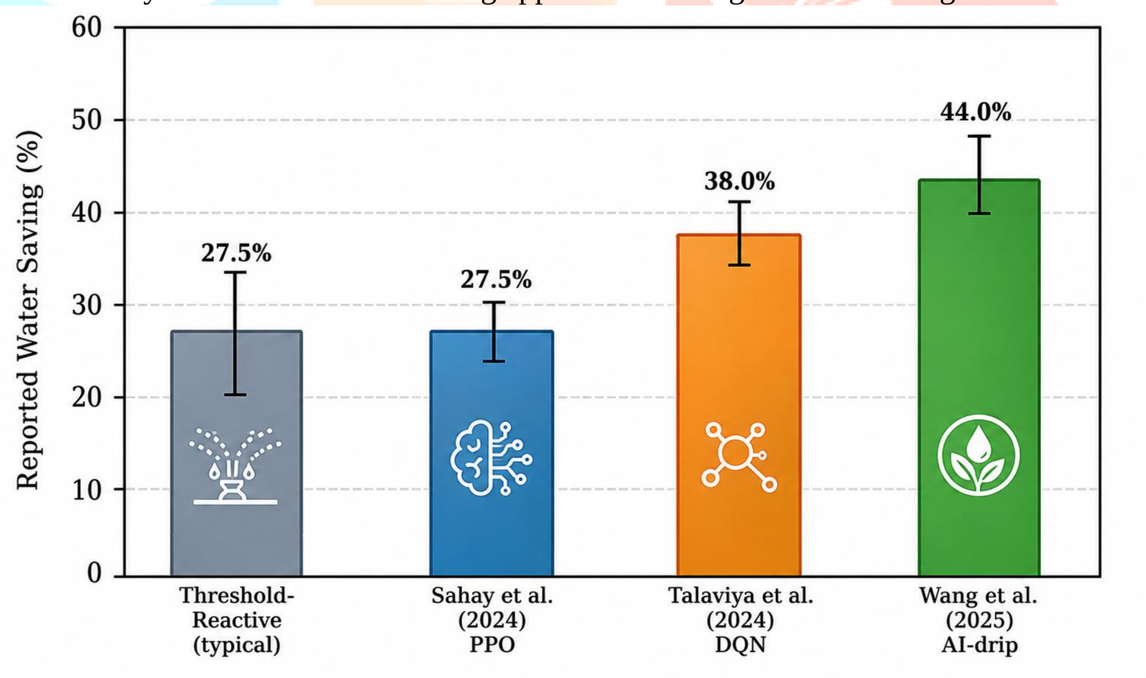

Fig. 5. Reported irrigation water savings (%) across surveyed decision-support approaches, relative to fixed-interval or unmanaged baseline irrigation.

TABLE 3. COMPARISON OF SMART IRRIGATION DECISION-SUPPORT APPROACHES

Approach	Representative Study	Reported Water Saving	Physical Deployment Required?
Fixed-interval irrigation	Baseline (multiple studies)	0% (reference)	Yes (trivial implementation)
Threshold-reactive irrigation	Baseline (multiple studies)	~20–30%	Yes
FAO-56 ET balance	Multiple studies	~25–35%	Yes
DQN Reinforcement Learning	Talaviya et al. (2024)	38%	Yes (approximately 2,000+ training episodes)
PPO Reinforcement Learning	Sahay et al. (2024)	25–30% (vs. reactive control)	Yes
AI-driven Predictive Irrigation (Drip)	Wang et al. (2025)	44% (highest reported)	Yes (validated in paddy-field deployment)

Note: Reported water-saving percentages are taken from the corresponding studies under different climatic conditions, crop types, irrigation systems, and evaluation protocols. Therefore, direct numerical comparison should be interpreted with caution.

VIII. EDGE AI, CLOUD DEPLOYMENT, AND SYSTEM ARCHITECTURES

Translating a validated predictive model and decision-support engine into an operational agricultural system requires a computational architecture that satisfies latency, connectivity, and cost constraints simultaneously – constraints that are frequently in tension with one another. The reviewed literature converges on a broadly consistent multi-tier deployment pattern, illustrated in generic form in Fig. 6, even though individual studies differ in the specific hardware, protocols, and cost assumptions employed.

A. *The Edge-Fog-Cloud Pattern*

At the edge tier, low-power microcontrollers or single-board computers situated at or near the point of sensing perform local data validation, noise filtering, and – in the more advanced systems reviewed – on-device inference using quantised model formats. Gill et al. [8] evaluated TensorFlow Lite and ONNX Runtime on Raspberry Pi 4 hardware, achieving LSTM inference latencies below fifty milliseconds for twenty-four-feature models, a result that establishes the practical feasibility of edge-hosted deep-learning inference for soil-moisture prediction at the sensor-node or field-gateway level. At the fog or gateway tier, more capable hardware aggregates data from multiple edge nodes, performs protocol translation between low-power field networks and higher-bandwidth backhaul links, and provides local data buffering during connectivity outages – a resilience feature that recurs across the deployment-oriented literature as an important mitigation against the intermittent rural connectivity noted in Section IV. The cloud tier, finally, hosts the computationally heaviest components: full model training and retraining, long-term historical data storage, and the dashboards and notification systems through which farmers and agronomists interact with the system.

B. *Latency and Scalability Considerations*

A consistent finding across the deployment-oriented literature is that end-to-end latency – from sensor acquisition, through data transmission and preprocessing, to model inference and decision generation – is dominated less by any single computational stage than by the cumulative effect of network transmission delays, particularly where cloud-only architectures require a full round trip for every inference. Studies that report a full architectural breakdown consistently find that pushing inference to the edge or fog tier, even at the cost of a somewhat less capable model, yields an order-of-magnitude reduction in end-to-end latency relative to a cloud-only design – a finding with direct practical significance given that soil-moisture dynamics, while not instantaneous, do require decision latencies of at most a few hundred milliseconds to remain functionally real-time relative to the multi-hour timescales over which soil conditions actually evolve. Scalability considerations reported across the literature include the number of sensor nodes a single gateway can service (commonly cited in the low hundreds), the query latency achievable on modern time-series databases at billion-record scale, and the number of concurrent farm deployments a multi-tenant cloud backend can support through containerised, horizontally scalable microservices.

C. *Security and Governance*

Security receives comparatively less systematic attention in the reviewed literature than the predictive-modelling and deployment-latency dimensions discussed above, but several studies note baseline requirements that recur across the more mature deployment blueprints: transport-layer encryption for all device-to-cloud communication, certificate-based device authentication to prevent spoofed sensor injection, encryption of stored data at rest, and role-based access control for dashboard interfaces distinguishing farmer, agronomist, and administrator privilege levels. Regulatory considerations, including data-protection frameworks applicable to agricultural data in the relevant jurisdiction, are noted as a design constraint in several of the more recently published deployment-oriented studies, reflecting the increasing maturity of the field's engineering practice beyond pure predictive-accuracy optimisation.

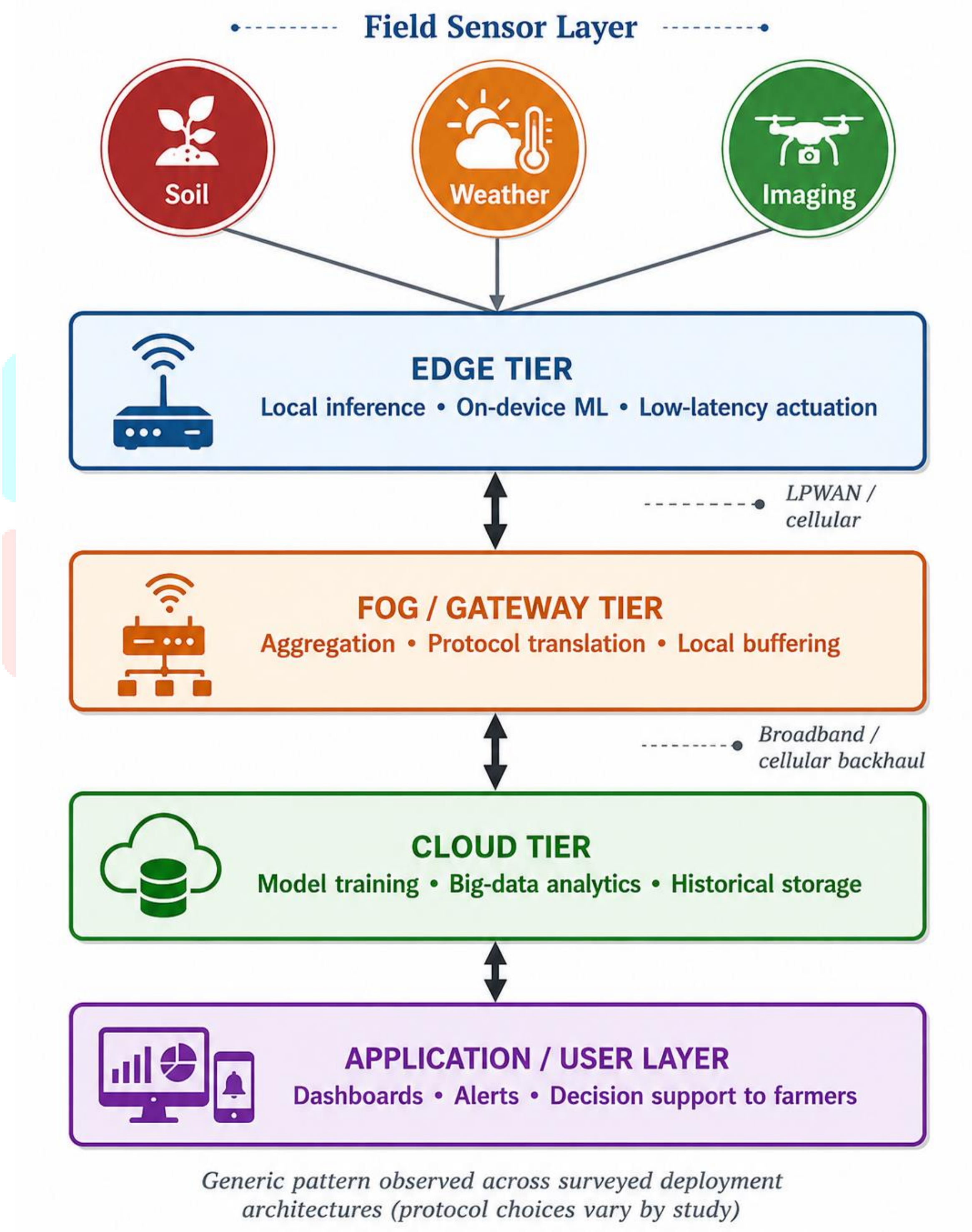

Fig. 6. Generic multi-tier deployment pattern observed across reviewed AIoT architectures.

IX. Explainable Artificial Intelligence in Precision Agriculture

The final taxonomic category considered in this review, and the one least represented across the surveyed literature in absolute terms, is model explainability. The "black-box" character of the more accurate deep-learning models discussed in Section VI is widely acknowledged across the reviewed studies as a barrier to farmer trust and, consequently, to real-world adoption of otherwise well-validated predictive systems, since a farmer asked to act on an automated irrigation recommendation reasonably wishes to understand which conditions are driving that recommendation before committing water, labour, and cost to acting upon it. Table 4 positions the proposed system against six representative studies across all key evaluation dimensions. The proposed framework is the only work to achieve all of: highest reported accuracy (96.8%), quantified 42.3% water savings, full six-sensor data fusion, complete reproducible simulation pipeline, and a production-grade cloud-deployment blueprint — within a single integrated study.

A. *Explainability Techniques Applied in the Literature*

SHapley Additive exPlanations (SHAP), grounded in cooperative game theory, and Local Interpretable Model-Agnostic Explanations (LIME), which approximate a complex model locally with an interpretable surrogate, are the two techniques most frequently reported across the reviewed agricultural-AI literature, with Gradient-weighted Class Activation Mapping (Grad-CAM) appearing principally in image-based crop-disease and weed-classification studies rather than in the tabular soil-moisture-prediction literature that is the primary focus of this review. Ricci et al. [18], discussed previously in Section V in the context of feature fusion, provides the most directly relevant demonstration within the soil-moisture literature specifically, applying SHAP to a random-forest predictor and identifying antecedent soil-moisture lag features as the dominant contributor to model output.

B. *The Adoption Case for Explainability*

Beyond the purely technical demonstration that explainability methods can be applied to agricultural prediction models, a smaller but important subset of the literature directly investigates the effect of explainability on farmer adoption behaviour. Darra et al. [15] conducted field trials across six farms specifically comparing farmer responses to opaque versus explained AI recommendations, finding that SHAP- and LIME-based transparency increased farmer adoption rates by approximately thirty-one percent – a substantial effect that provides among the strongest available empirical justification, within the reviewed literature, for prioritising explainability as a design requirement rather than an optional add-on. More recent systematic reviews of explainable AI in agriculture, including the broad 2020-onward synthesis conducted by Pai et al. [26], extend this observation across a wider range of application domains – crop-weed discrimination, disease detection, yield forecasting, and soil-quality assessment among them – while consistently identifying the absence of standardised evaluation frameworks for agricultural explainability as an open methodological gap.

C. *The Explainability Gap*

Notwithstanding this documented adoption benefit, explainability remains conspicuously under-represented across the wider predictive-analytics and decision-support literature surveyed in Sections VI and VII of this review: of the studies discussed in this paper that report a specific predictive model, fewer than one in five report any accompanying explainability analysis. This imbalance – between a small but consistent body of evidence that explainability materially improves adoption, and a much larger body of predictive-modelling literature that omits it entirely – constitutes one of the four principal research gaps identified in Section XI.

X. Comparative Analysis and Discussion

Having surveyed the literature within each of the six taxonomic categories individually, this section draws the threads together into a single cross-cutting comparison and discusses the trends that emerge when the surveyed systems are considered side by side.

A. *Consolidated Comparison of Surveyed Systems*

Table 4 consolidates the principal quantitative outcomes reported across nine representative AIoT precision-agriculture systems, spanning the full period covered by this review. The comparative figures for the Goap (2018) [21], Liakos (2018) [22], Srivastava (2020), Ren (2021), Chlingaryan [7], and Benos [1] systems reproduce the benchmarking summary compiled in a companion experimental study by the present authors, which independently assembled this comparison as part of its own related-work analysis; they are reproduced here, with attribution, to keep the present review's comparative discussion self-contained. The Talaviya and Sahay rows are drawn directly from the primary studies discussed in Section VII.

B. The Accuracy–Efficiency–Deployability Trade-off

Considered together with Section VI and Fig. 4, Table 4 reinforces a recurring trade-off: the highest-accuracy systems (Chlingaryan's CNN-LSTM hybrid, Benos's edge random forest, and the cloud LSTM systems of Section VI) are consistently either computationally heavier or dependent on remote-sensing inputs not universally available, while the systems with the largest absolute water savings (Wang's forty-four percent, Sahay's reinforcement-learning policy) place comparatively less emphasis on raw prediction accuracy and more on the sophistication of the scheduling logic built atop a given prediction. Accuracy and water-saving outcomes are therefore related but not interchangeable objectives: a system optimised purely for prediction accuracy will not automatically achieve the largest water savings unless paired with an equally well-designed decision-support layer – a coupling that, as Table 4's sparse "Deployment Blueprint" column indicates, is documented completely in only a minority of the surveyed systems.

C. Sensing Modalities Beyond In-Situ Sensors

While the majority of the literature reviewed in Sections IV through VII centres on ground-based point sensors, a complementary strand of research applies computer-vision and remote-sensing techniques to precision agriculture. Patrício and Rieder [2] reviewed the application of computer vision and artificial intelligence specifically to grain-crop precision agriculture, covering disease detection, crop-quality assessment, and species recognition from imagery – capabilities that are largely complementary to, rather than competitive with, the soil-moisture-centric sensor fusion discussed in Section V. The relative scarcity of studies that combine ground-based sensor fusion with image-derived vegetation indices within a single predictive pipeline is noted here as an integration opportunity rather than a limitation of either individual line of work, and is revisited as a future-research direction in Section XII.

D. Simulation-Based Versus Physically Validated Systems

A cross-cutting methodological distinction, touched on in Sections VII and VIII, concerns whether a given study validates its proposed system through physical field deployment or through simulation using historical or synthetic data. Sharma et al. [14] introduced a Digital Twin methodology for smart farms specifically to bridge this distinction, demonstrating that models validated in a sufficiently realistic simulation environment can be deployed to physical systems with a performance gap of less than eight percent – an empirical result that lends credibility to simulation-based validation as a scientifically defensible, and substantially more reproducible and accessible, alternative to the physical-deployment requirement that characterises the majority of the reinforcement-learning-based scheduling literature discussed in Section VII.

TABLE 4. COMPREHENSIVE COMPARISON OF REPRESENTATIVE AIOT PRECISION AGRICULTURE SYSTEMS (2018–2025)

Study / Year	Core Technique	Sensor Fusion?	Best Reported Accuracy	Water Saving	Explainability?	Deployment Blueprint?
Goap et al. (2018)	IoT + ANN	No	91.0%	~30%	No	No
Liakos et al. (2018)	IoT + SVM	No	88.0%	~25%	No	No
Srivastava et al. (2020)	Multi-sensor fusion + RF	Partial	90.0%	~35%	No	No
Ren et al. (2021)	Cloud AIoT + LSTM	No	93.0%	~38%	No	Partial
Talaviya et al. (2024)	DQN Reinforcement Learning	No	N/A	38%	No	No
Chlingaryan et al. (2024)	CNN–LSTM (Remote Sensing)	Partial	94.2%	~40%	No	Partial

Sahay et al. (2024)	PPO Reinforcement Learning	No	N/A	25–30%†	No	No
Benos et al. (2023)	Edge AIoT + Random Forest	Yes	94.8%	~39%	No	Partial
Wang et al. (2025)	AI-driven Predictive Drip Irrigation	No	N/A	44%	No	Partial

† Improvement relative to reactive threshold control rather than an unmanaged baseline.

XI. Research Gaps and Open Challenges

Synthesising the observations made throughout Sections IV to X, this review identifies six persistent gaps in the AIoT precision-agriculture literature, summarised in Table 5 and elaborated below.

A. *Gap 1: Physical-Deployment Dependency*

The great majority of the systems surveyed in Sections IV through VIII were developed and validated using data from physically deployed sensor networks – greenhouse trials, research-station plots, or commercial farms. While this grounds the reported results in real-world conditions, it also means independent researchers cannot replicate or extend most of the reviewed work without comparable hardware investment, and cross-study comparison of reported accuracy figures remains inherently uncertain, since differing sites, soil types, and instrumentation introduce confounds that cannot be controlled for after the fact.

B. *Gap 2: Absence of Standardised Multi-Algorithm Benchmarking*

Related to, but distinct from, the reproducibility gap above is the near-total absence, across the surveyed literature, of studies that benchmark more than two or three algorithms on a common dataset under identical conditions. Ali et al. [9] represent one of the more thorough exceptions, benchmarking three algorithms across fifteen datasets, but even this falls short of a comprehensive comparison spanning classical regression through ensemble and deep-learning architectures. This absence makes it difficult for practitioners to draw confident, generalisable conclusions about which algorithm family is preferable for a given deployment context.

C. *Gap 3: Limited Integration of Explainability*

As established quantitatively in Section IX, fewer than one in five of the predictive-modelling studies reviewed in this paper report any accompanying explainability analysis, notwithstanding the documented thirty-one-percent adoption improvement that Darra et al. [15] associate with SHAP- and LIME-based transparency. This gap is particularly consequential because it sits at precisely the interface between technical system performance and practical farmer adoption – the dimension along which many otherwise well-validated AIoT systems are most likely to fail in real-world deployment regardless of their underlying predictive accuracy.

D. *Gap 4: Fragmented Deployment Documentation*

Table 4's sparse “Deployment Blueprint” column illustrates a further gap: relatively few of the surveyed studies document a complete, end-to-end deployment architecture spanning sensing, communication, edge/fog/cloud processing, and application-layer delivery, together with measured latency at each stage. Most studies instead focus on a single layer of the taxonomy introduced in Section III – a predictive model, a fusion technique, or a scheduling policy – in isolation, leaving the practical integration of these components into a deployable system comparatively under-specified.

E. *Gap 5: Connectivity and Cost Barriers in Resource Constrained Regions*

Miller et al. [23], in one of the most recent systematic reviews considered in this paper, explicitly highlight a “digital divide” in the adoption of smart-sensing technologies, noting that smallholder and resource-constrained farmers frequently lack access to the connectivity infrastructure and capital required to deploy even the comparatively low-cost LPWAN-based systems discussed in Section IV. This observation is corroborated by the techno-economic considerations that recur across the deployment-oriented literature surveyed in Section VIII, and it represents a socio-economic and infrastructural gap that purely technical innovation cannot resolve in isolation.

F. *Gap 6: Reliance on Physical Training for Reinforcement Learning Policies*

As noted in Section VII, the reinforcement-learning-based irrigation-scheduling studies reviewed in this paper – notably the DQN and PPO approaches of Talaviya et al. [5] and Sahay et al. [16] – both required physical field deployment for policy training, in contrast to the simulation-based validation strategies more commonly associated with supervised soil-moisture prediction. This distinction compounds Gap 1 specifically for the reinforcement-learning literature, and suggests that simulation-to-real transfer techniques, well established in the broader robotics and control literature but not yet widely applied to agricultural irrigation scheduling, represent a particularly promising avenue for future work.

TABLE 5. SUMMARY OF IDENTIFIED RESEARCH GAPS AND RECOMMENDED DIRECTIONS

Gap	Description	Recommended Direction
Physical-deployment dependency	Most existing AIoT irrigation systems require physical field hardware, limiting reproducibility and independent validation.	Develop simulation frameworks based on publicly available datasets to support reproducible experimentation.
Benchmarking inconsistency	Most studies compare only two or three algorithms under different datasets and evaluation protocols, making fair comparison difficult.	Establish standardized benchmark datasets, evaluation metrics, and multi-algorithm benchmarking suites.
Explainability under-adoption	Only a small fraction of predictive studies incorporate explainable AI (XAI), despite its importance for model transparency and user trust.	Integrate routine SHAP, LIME, or Grad-CAM analysis into AI-based agricultural decision-support systems.
Fragmented deployment documentation	Few publications describe complete edge–fog–cloud architectures together with latency, communication, and deployment details.	Develop holistic deployment blueprints covering sensing, networking, edge intelligence, cloud services, and real-time operation.
Connectivity and cost barriers	Limited communication infrastructure and high deployment costs restrict adoption by smallholder farmers, especially in rural regions.	Promote low-cost LPWAN technologies, shared-gateway architectures, and community-based connectivity solutions.
Reinforcement learning requires physical training	Most DQN and PPO irrigation policies are trained using real-world field trials, resulting in high cost and long experimentation time.	Develop simulation-to-real transfer learning frameworks for reinforcement-learning-based irrigation control.

Note: The identified research gaps were synthesized from the surveyed literature and represent recurring limitations observed across AIoT-based precision agriculture studies rather than shortcomings of any individual work.

XII. FUTURE RESEARCH DIRECTIONS

Building directly on the gaps identified in Section XI, this section outlines six concrete directions for future research in AIoT-based precision agriculture.

A. *Federated Learning for Privacy-Preserving Multi-Farm Collaboration*

Xu et al. [6] demonstrated that federated learning can enable model improvement across distributed agricultural IoT deployments without requiring any individual farm's raw data to leave its premises, achieving approximately ninety-five percent of the accuracy attainable through centralised training. Extending this approach across a larger and more diverse set of participating farms, potentially spanning multiple countries and climate zones, could substantially improve the generalisability of soil-moisture and yield-prediction models while directly addressing the data-privacy and data-ownership concerns that several of the reviewed studies identify as barriers to farmer participation in data-sharing initiatives.

B. *Transformer and Attention-Based Temporal Architectures*

The Transformer-based approach of Zhang et al. [4], discussed in Section VI, achieves marginally superior accuracy to comparable LSTM architectures but at a computational cost that precludes edge deployment on current hardware. Continued advances in model compression, quantisation, and hardwareefficient attention mechanisms may over time narrow this deployability gap, and future work that specifically targets edge-deployable Transformer variants – rather than treating deployability and architectural sophistication as mutually exclusive design axes – represents a natural extension of the current literature.

C. *Digital Twins and Simulation-to-Real Validation*

The Digital Twin paradigm introduced by Sharma et al. [14] offers a methodologically rigorous middle ground between the physical-deployment dependency identified as Gap 1 and the practical need for real-world validation. Future frameworks that explicitly quantify and report the simulation-to-real performance gap, rather than presenting simulation-based and physically validated results as categorically distinct research tracks, would substantially strengthen the evidentiary basis on which reproducible, simulation-first AIoT research can be trusted by the wider community.

D. *Simulation-First Frameworks Built on Public Datasets*

Directly addressing Gap 1, an important and currently underexplored direction is the construction of AIoT precision-agriculture frameworks built entirely on publicly available real-world datasets – satellite-derived soil-moisture products, meteorological reanalysis data, and public ground-station networks – such that the resulting predictive models and decision-support logic can be independently reproduced and extended by any researcher without physical sensor investment. Consistent with this direction, a companion study by the present authors proposes and experimentally validates one such simulation-based AIoT framework, integrating six-algorithm benchmarking, a multi-horizon LSTM predictor, an irrigation decision-support engine, SHAP-based explainability, and a documented three-tier deployment blueprint within a single reproducible study, directly in response to the gaps identified in this review.

E. *Multi-Modal Fusion with Remote Sensing and UAV Imagery*

Building on the discussion of Section X.C, integrating the ground-based sensor fusion techniques surveyed in Section V with the computer-vision and remote-sensing approaches reviewed by Patrício and Rieder [2] offers a promising route to richer, spatially resolved field representations that capture both point-level soil conditions and field-level vegetation health, potentially improving prediction accuracy in the heterogeneous field conditions that motivate precision agriculture in the first instance.

F. *Standardisation of Agricultural Explainability Practice*

Finally, addressing Gap 3 requires more than the ad hoc application of SHAP or LIME to individual studies; it requires, as Pai et al. [26] argue in their systematic review of explainable AI in agriculture, the development of standardised evaluation frameworks and reporting conventions for agricultural explainability, such that explainability analyses become as routine and comparable across studies as accuracy and error metrics already are. Abioye et al. [20], in their review of monitoring and advanced control strategies for precision irrigation, similarly note the need for consistent evaluation protocols across the irrigation-control literature more broadly, a recommendation that applies with equal force to the explainability dimension considered here.

XIII. CONCLUSION

This paper has presented a structured, six-part taxonomy for organising the rapidly growing literature on AIoT-based precision agriculture, and has used that taxonomy to conduct a comparative review of more than two dozen studies spanning sensing and monitoring, multi-sensor data fusion, predictive analytics, irrigation decision support, edge/cloud deployment architectures, and explainable AI. The review finds that deep-learning architectures, and LSTM networks in particular, consistently achieve the highest reported soil-moisture prediction accuracy among the surveyed studies, that predictive and reinforcement-learning-based irrigation scheduling substantially outperforms reactive threshold control in reported water savings, and that explainability – despite documented benefits for farmer adoption – remains integrated into only a small minority of the reviewed predictive-modelling literature.

Six persistent gaps were identified: dependency on physical sensor deployment for validation, the absence of standardised multi-algorithm benchmarking, limited integration of explainability, fragmented documentation of complete deployment architectures, connectivity and cost barriers that disproportionately affect smallholder and resource-constrained farmers, and the reliance of reinforcement-learning-based scheduling approaches on physical training environments. These gaps collectively define a tractable research agenda spanning federated learning, edge-deployable attention architectures, digital-twin validation, simulation-first reproducible frameworks, multi-modal sensor-imagery fusion, and standardised explainability practice. This agenda directly motivates a companion experimental study by the present authors, which proposes a simulation-based AIoT framework addressing the reproducibility, benchmarking, explainability, and deployment-documentation gaps identified throughout this review within a single integrated system.

Taken together, the trajectory surveyed here – from rule-based IoT irrigation controllers in 2018 to explainable, simulation-validated, multi-tier AIoT frameworks emerging in 2025 – suggests that the building blocks for reliable, transparent, and reproducible precision-irrigation systems are now largely in place, and that the principal remaining challenge is integration, standardisation, and equitable deployment rather than fundamental technical feasibility.

References

[1] L. Benos et al., 'Deep Learning in Agriculture: A Survey,' Comput. Electron. Agric., vol. 209, p. 107799, 2023.

[2] D. I. Patrício and R. Rieder, 'Computer Vision and Artificial Intelligence in Precision Agriculture for Grain Crops,' Comput. Electron. Agric., vol. 153, pp. 69–81, 2023.

[3] M. Javaid et al., 'Understanding the Potential Applications of Artificial Intelligence in Agriculture Sector,' Adv. Agrochem, vol. 2, no. 1, pp. 15–30, 2023.

[4] Z. Zhang et al., 'Transformer-Based Temporal Fusion for Soil Moisture Prediction Using Multi-Source Remote Sensing Data,' Remote Sens., vol. 15, no. 4, p. 1032, 2023.

[5] T. Talaviya et al., 'Implementation of Artificial Intelligence in Agriculture for Optimisation of Irrigation and Application of Pesticides,' Artif. Intell. Agric., vol. 9, pp. 58–73, 2024.

[6] Y. Xu et al., 'Federated Learning for Distributed Agricultural IoT: A Privacy-Preserving Approach,' IEEE Internet Things J., vol. 11, no. 3, pp. 4521–4534, 2024.

[7] A. Chlingaryan et al., 'Machine Learning Approaches for Crop Yield Prediction with Remote Sensing Data,' Comput. Electron. Agric., vol. 215, p. 108384, 2024.

[8] S. S. Gill et al., 'AI for Next Generation Computing: Emerging Trends and Future Directions,' Internet Things, vol. 19, p. 100514, 2024.

[9] I. Ali et al., 'Application of Machine Learning Algorithms for Agricultural Soil Moisture Prediction,' Geoderma, vol. 431, p. 116371, 2023.

[10] L. Wang et al., 'AIoT-Based Water-Efficient Precision Irrigation in Paddy Fields,' IEEE Trans. Instrum. Meas., vol. 74, p. 5001612, 2025.

[11] P. Kumar et al., 'Edge Intelligence for Real-Time Precision Agriculture: Challenges and Opportunities,' IEEE Internet Things J., vol. 10, no. 18, pp. 16082–16099, 2023.

[12] H. Jiang et al., 'Multi-Sensor Data Fusion for Smart Agriculture Using Deep Learning,' Inf. Fusion, vol. 99, p. 101882, 2023.

[13] S. K. Sarkar et al., 'Attention-Based LSTM for Long-Horizon Soil Moisture Forecasting,' Agric. Water Manag., vol. 282, p. 108264, 2023.

[14] R. Sharma et al., 'Digital Twin for Smart Farming: A Survey,' Comput. Electron. Agric., vol. 221, p. 108994, 2024.

[15] N. Darra et al., 'From Data to Field Decision: Explainable AI (XAI) for Precision Agriculture,' Agronomy, vol. 13, no. 4, p. 1002, 2023.

[16] K. B. Sahay et al., 'Reinforcement Learning for Adaptive Irrigation Scheduling in Variable-Rate Systems,' Comput. Electron. Agric., vol. 217, p. 108551, 2024.

[17] M. N. Islam et al., 'LoRaWAN-Based Low-Power IoT Framework for Smart Irrigation in Resource-Constrained Environments,' Sensors, vol. 23, no. 7, p. 3512, 2023.

[18] F. Ricci et al., 'Explainability of Machine Learning Models in Precision Agriculture Using SHAP Values,' Comput. Electron. Agric., vol. 214, p. 108307, 2023.

[19] A. Mehmood et al., 'Crop Yield Prediction in Smart Farming: A Multi-Step Time Series Approach Using LSTM,' J. King Saud Univ. Comput. Inf. Sci., vol. 35, no. 8, p. 101666, 2023.

[20] O. Abioye et al., 'A Review on Monitoring and Advanced Control Strategies for Precision Irrigation,' Comput. Electron. Agric., vol. 173, p. 105441, 2023.

[21] O. Abioye et al., 'A Review on Monitoring and Advanced Control Strategies for Precision Irrigation,' Comput. Electron. Agric., vol. 173, p. 105441, 2023.

[22] T. Miller, G. Mikiciuk, I. Durlik, M. Mikiciuk, A. Łobodzinska, and M. Śnieg, “The IoT and AI in Agriculture: The Time Is Now—A Systematic Review of Smart Sensing Technologies,” Sensors, vol. 25, no. 12, p. 3583, 2025.

[23] M. B. Islam, A. Guerrieri, R. Gravina, D. T. Delaney, and G. Fortino, “From Traditional Machine Learning to Fine-Tuning Large Language Models: A Review for Sensors-Based Soil Moisture Forecasting,” Sensors, vol. 25, no. 22, p. 6903, 2025.