signals

MDPI

Article

Sensor Data Transmission in Smart Agriculture: Challenges, Potentials, and Limitations in the Fourth Industrial Revolution

Alvertos Ioannis Mourikis, Iosif Polenakis *, Christos Sarantidis, Romanos Kalamatianos and Ioannis Karydis

Department of Informatics, Ionian University, 49100 Corfu, Greece; amourikis@ionio.gr (A.I.M.); csarantidis@ionio.gr (C.S.); rkalam@ionio.gr (R.K.); karydis@ionio.gr (I.K.)
* Correspondence: ipolenakis@ionio.gr

Abstract

Smart agriculture is a key application domain of the Fourth Industrial Revolution (4IR), where cyber–physical systems, the Internet of Things (IoT), and data-driven analytics converge to improve productivity, sustainability, and resource efficiency. In this context, large volumes of sensor data collected in the field form the basis for automated monitoring, decision support, and adaptive control across agricultural processes. However, the value of such data depends not only on sensing technologies and connectivity, but also on how data are generated, transferred, processed, and interpreted throughout their lifecycle. This paper is organized in two parts. The first part reports a narrative review of the literature, based on a documented search protocol, and provides an overview of sensor data and sensor data transmission networks in smart agriculture. It examines the main challenges, potentials, and limitations associated with IoT-based field deployments in 4IR environments, and consolidates them into ten open problems. These problems are ordered from the device layer to the socio-economic layer and span physical robustness, device security, rural connectivity, aerial sensing, interoperability, labelled-data scarcity, explainability, federated access control, data ownership, and cost. The second part responds to these open problems with the conceptualization of a layered architecture and a set of design proposals, covering hierarchical edge security, offline-first synchronization, semantic interoperability, explainable decision support, and farmer-centric data governance, that are grounded in established techniques and standards but have not been implemented or empirically validated in this work; they are intended as a design blueprint and a starting point for future prototyping and evaluation rather than as measured results. The paper is accordingly positioned as a conceptual, perspective-oriented contribution: it states the procedure by which the open problems were derived from the reviewed works, it states what the proposed integration is claimed to add over existing smart-agriculture architectures, and it sets out an evaluation plan that fixes the metrics and the baselines against which each proposed component would have to be tested. Throughout both parts, particular attention is given to the role of data integrity and provenance as foundational elements for trustworthy agricultural data, enabling reliable analysis, auditability, and long-term reuse. By synthesizing current approaches, open issues, and conceptual solution patterns, this paper aims to support the design of resilient, transparent, and future-ready smart agriculture systems aligned with the principles of the 4IR.

Keywords: smart agriculture; agricultural Internet of Things; sensor data transmission networks; data integrity and provenance; edge computing; data governance

Check for updates

Academic Editor: Ying-Ren Chien

Received: 3 July 2026
Revised: 22 September 2026
Accepted: 30 September 2026
Published: 10 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Smart agriculture in the 4IR integrates IoT, Artificial Intelligence (AI), cloud computing, and advanced sensors [1,2]. Together, these technologies enable data-driven farming through the real-time monitoring of soil, climate, and crop conditions. They are widely reported to improve productivity, resource efficiency, and sustainability. These technologies support precision agriculture practices such as automated irrigation, fertilization, and pest control, often enhanced by drones and autonomous systems [3,4]. However, the increasing reliance on interconnected sensor networks also introduces significant risks. Data manipulation, unauthorized access, and cyberattacks can disrupt agricultural operations and distort the decisions taken on the basis of the collected data [5–7]. Ensuring sensor security through encryption, secure authentication, intrusion detection, and resilient edge computing is essential to maintain data integrity and protect smart farming systems [8–10].

This work investigates smart agriculture within the 4IR. It focuses on the sensor data produced by field-based IoT deployments and on the transmission networks that carry that data from the field to the point of use.The central argument is that trust, integrity, and traceability are core enabling factors for reliable decision-making. To keep the introduction focused, the review of the literature is presented separately in Section 2, and the resulting ten open problems are stated in Section 4. Sections 5 and 6 then respond to those problems with a conceptual layered architecture and a set of design proposals. These are intended as a blueprint for future implementation and evaluation rather than as validated engineering results.

The positioning of the paper follows from that division of labour. The first part is a review; the second is a design proposal. Neither part reports a prototype, a simulation, a field trial, or a user study, and the paper therefore makes no empirical claim about the robustness, the effectiveness, or the economic benefit of the architecture it describes. It is offered as a conceptual, perspective-oriented contribution with three components: a consolidation of what the reviewed literature leaves unresolved, a set of design patterns that plausibly address those gaps, and an evaluation plan by which the patterns can later be tested. Section 6.3 states what the integration is claimed to add over existing smart-agriculture architectures, and Section 6.4 states the evaluation plan.

1.1. Trust, Integrity, and Traceability as Enabling Factors

In the context of Smart Agriculture and the 4IR, sensor data constitute a fundamental basis for monitoring, analysis, and decision-making in agricultural systems [1]. Measurements originating from soil sensors, weather stations, cameras, traps, drones, and IoT devices are used for irrigation, fertilization, disease detection, risk assessment, and production optimization [2]. However, the value of these data depends primarily on the extent to which they can be considered trustworthy, accurate, and interpretable [11].

For this reason, trustworthiness is treated in this work as a property of the system as a whole. Protection against unauthorised access or attack is only one aspect of the problem. Data in an agricultural information system can be inaccurate for many other reasons, including sensor malfunction, incorrect calibration, packet loss, ambient noise, poor placement of measurement devices, and deliberate manipulation [5]. The consequences are operational rather than merely informational. Incorrect soil moisture readings may trigger excessive or insufficient irrigation. Incorrect temperature and humidity readings may distort disease prediction. Manipulated data from traps and cameras may trigger unnecessary or misdirected crop protection measures [3,12]. In a cyber–physical agricultural setting, such errors translate into financial losses, environmental impact, and yield losses [8].

Data integrity is therefore a precondition for any further use of the data. Automation, artificial intelligence, and decision support all require reasonable assurance that data are

accurate, complete, and consistent before they are used for prediction [8]. Integrity also involves protection against malicious activities as well as the ability to identify errors, deviations, quality degradation and inconsistencies when collecting and transmitting, storing and processing data [9].

Data provenance is an equally important and complementary concept. Proper interpretation, auditing, and reuse of a measurement require knowing which sensor produced it, when, under which conditions, with which calibration, and through which transmission and processing steps [5,8,9,11]. Without provenance information, a measurement may be difficult to assess, compare with other measurements or use in a different situation. Provenance provides benefits in terms of traceability, documentation, quality control, error analysis and ethical data use in Smart Agriculture [8,9,11].

1.2. Contributions and Structure of the Paper

This paper intends to provide a conceptual overview of the potential, challenges and constraints of sensor data in Smart Agriculture, with a focus on the key characteristic of trustworthiness as a system property. This paper contributes in three ways:

1. It identifies data integrity, data quality, and provenance as key factors for trustworthy decision-making, and consolidates the reviewed literature into ten open problems.
2. It describes mechanisms and design patterns with the potential to enable trustworthy data collection, transmission, storage, and use.
3. It explores design considerations for constrained environments, in which sensors, edge devices, and agricultural IoT devices have limited energy, processing, communication, and maintenance capabilities.

By doing so, this study is not limited to a very technical debate on cybersecurity. It looks more generally at how agricultural systems can be structured so that they generate, handle, and use data that are trusted, auditable, and appropriate for responsible decision-making.

The remainder of the paper is organised as follows. Section 2 reviews the related literature and states the search protocol on which the review rests. Section 3 presents the conceptual foundations of smart-agriculture cyber–physical systems, the enabling technologies, the trade-offs between sensor data transmission networks, and the sensor data lifecycle. Section 4 states the ten open problems that follow from the review. Section 5 presents a conceptual solution for each of them, in the same order. Section 6 integrates those solutions into a single layered architecture. Section 7 concludes and identifies directions for future work.

2. Literature Review

This section reviews the literature on sensor data and sensor data transmission networks in smart agriculture. Section 2.1 states the protocol by which the reviewed works were identified. Sections 2.2–2.9 then present the reviewed works in eight thematic clusters. The clusters proceed from the technical layers of an agricultural sensing system towards its policy, governance, and adoption aspects: sensing hardware, transmission networks, edge and architecture, aerial platforms, analytics, field applications, traceability, and finally security, privacy, and governance. Section 2.10 synthesises the points of consensus and disagreement that emerge across the clusters.

2.1. Review Scope and Search Protocol

This paper reports a narrative rather than a systematic review, and the protocol below is stated so that the basis of the selection is transparent and repeatable. Records were retrieved from Scopus, IEEE Xplore, ScienceDirect, SpringerLink, and MDPI. Google

Scholar was used for backward and forward citation chasing from the works retained at full-text stage. The search combined three groups of terms with Boolean operators. The first group covered the domain ("smart agriculture", "precision agriculture", "Agriculture 4.0", "digital farming"). The second covered the technology ("IoT", "wireless sensor network", "sensor data", "LPWAN", "LoRaWAN", "NB-IoT", "edge computing", "UAV"). The third covered the system property of interest ("security", "interoperability", "data integrity", "provenance", "explainability", "data governance", "adoption"). The search covered publications from January 2016 to June 2025, with the lower bound chosen because the 4IR and LPWAN literature in agriculture consolidates from that point onwards.

Approximately 210 records remained after de-duplication. Titles and abstracts were screened against three inclusion criteria: the work had to address agricultural sensing, agricultural data transmission, or the analysis of agricultural sensor data; it had to be a peer-reviewed article, conference paper, or review, or a normative standards document; and it had to be written in English. Editorials, extended abstracts, and papers reporting only crop-science outcomes without a sensing or data component were excluded. Full-text assessment of the remaining records prioritised recent reviews and surveys with broad coverage, together with primary studies that report deployment experience, and resulted in the 55 works cited here.

Two limitations follow directly from this protocol and are restated in Section 7. First, the selection reflects the judgement of the authors rather than a pre-registered protocol, so the review is subject to selection bias and no claim of exhaustive coverage is made. Second, screening was performed by the author team without independent double-screening, so no inter-rater agreement can be reported.

2.2. Sensing Hardware and Wireless Sensor Networks

Aarif et al. [13] survey advances in smart sensor technology and report that miniaturisation and low-power design have made continuous monitoring of soil parameters and climatic variables practical at field scale. Yadav et al. [14] examine emerging frontiers in nanotechnology and find that nano-biosensors offer high sensitivity in detecting soil moisture and plant pathogens. Soussi et al. [15] review smart sensors and smart data for precision agriculture, and identify techniques such as Laser-Induced Breakdown Spectroscopy as useful for measuring soil nutrient levels. Turning from individual devices to networks, Ayaz et al. [16] argue that WSNs are central to field monitoring, and that multi-hop architectures are best suited to optimising irrigation and pest control over large areas. Mowla et al. [17] provide a systematic review of WSN protocols and compare ZigBee, Wi-Fi, and LoRaWAN with respect to energy efficiency.

2.3. Connectivity and Sensor Data Transmission

Pagano et al. [18] survey the use of LoRa in Agriculture 4.0 and report that long-range networks combined with data compression can extend sensor battery life substantially. Aldhaheri et al. [19] extend that analysis to the opportunities and open challenges of LoRa communication for Agriculture 4.0. Sanjeevi et al. [20] evaluate heterogeneous wireless networks, including Wi-Fi, WiMAX, and Long Term Evolution (LTE), using Fractional Order Darwinian Particle Swarm Optimization, and report reduced latency and increased throughput in agricultural data pipelines. Looking further ahead, Polymeni et al. [21] assess the convergence of IoT with 6G and suggest that terahertz spectral imaging and quantum sensing may support ultra-massive connectivity in future agricultural networks.

2.4. Edge Intelligence and IoT Architectures

Shi et al. [10] set out the general case for edge computing, arguing that processing close to the data source is necessary when bandwidth, latency, or privacy constrain a cloud-

centric design. In agriculture, Qazi et al. [22] critically assess AI-enhanced smart agriculture and find that deep learning on edge devices can identify threats in real time and support adaptive control. Mansoor et al. [23] examine the integration of smart sensors and IoT, and report that soil moisture sensors, such as those based on Time Domain Reflectometry, combined with edge computing, allow irrigation to be managed autonomously at the field. Kassim [24] notes that IoT implementation raises practical difficulties, and that open-source software and robust network design are needed to support remote monitoring. At the architectural level, Araújo et al. [25] review developments in agricultural IoT and propose a cloud-based IoT architecture for future smart agricultural systems. Adli et al. [26] define the Artificial Intelligence of Things (AIoT) and conclude that combining AI with IoT devices allows image processing and environmental prediction to be performed on the device itself.

2.5. Remote Sensing and UAV Platforms

Phang et al. [27] review the progression from satellite to UAV-based remote sensing, and report that support vector machines and recurrent neural networks can map crop types and field variability accurately. Khanal et al. [28] survey accomplishments and limitations in agricultural remote sensing, and find that although sensors recording visible and near-infrared data are mature, scaling hyperspectral platforms remains technically difficult. Guebsi et al. [29] report that drones carrying multispectral and hyperspectral cameras can compute vegetation indices that indicate plant vigour and stress, and argue that beyond visual line of sight (BVLOS) operation is necessary for precision agriculture at scale. Delavarpour et al. [30] present a technical study of UAV characteristics and show that careful geometric and radiometric correction is required to reduce weather effects and produce high-quality orthophotographs. Aslan et al. [31] review deep-learning-based UAV applications and report high accuracy for convolutional neural networks in weed detection and crop classification from aerial imagery. Maddikunta et al. [4] state that UAVs combined with collaborative computing frameworks overcome the spatial resolution constraints of conventional monitoring, while also documenting endurance, payload, and regulatory constraints.

2.6. Data Analytics, AI, and Decision Support

Bhat and Huang [32] assess the big data and AI revolution in precision agriculture and conclude that machine learning models can convert heterogeneous agricultural data into actionable knowledge. Garg and Alam [33] review the big data analytics literature systematically and find that deep learning and fuzzy logic dominate the optimisation of agricultural activities. Linaza et al. [34] study data-driven AI applications in European projects and report that robotic platforms using multi-view stereo algorithms can reconstruct canopy geometry in three dimensions; they also identify interpretability as a barrier to uptake. Zhai et al. [35] survey decision support systems for Agriculture 4.0 and discuss both the opportunities for data-driven crop management and the operational obstacles that remain. Vinod Chandra et al. [36] argue that the integration of AI and IoT underpins the transition to Agriculture 5.0, and demonstrate crop yield prediction under climatic fluctuation.

2.7. Precision Agriculture Applications in the Field

Karunathilake et al. [37] review the transition from traditional to smart farming and find that variable rate technologies and high-throughput phenotyping reduce input waste considerably. Sharma and Shivandu [38] review the integration of AI and IoT for crop monitoring and management, and report improvements in the timeliness and spatial resolution of crop health assessment. Alahmad et al. [39] compare site-specific seeding supported by multi-sensor data fusion with uniform application, and determine that variable rate application yields higher tuber yield and profitability. Monteiro et al. [40]

review precision agriculture instruments for crop and livestock farming and observe that non-invasive monitoring improves both resource efficiency and livestock welfare. Saiz-Rubio and Rovira-Más [41] outline an information-driven management cycle centred on the definition of management zones through proximal sensing, and conclude that managing in-field spatial variability is decisive. Hayat et al. [42] link biotechnology with precision agriculture and argue that genome editing combined with IoT monitoring can increase crop resilience sustainably. Sott et al. [43] conduct a bibliometric analysis of the coffee sector and find multispectral remote sensing to be the most prevalent technology for mapping spatial variability.

2.8. Traceability and Food Supply Chains

Chapman et al. [44] examine how the 4IR affects food supply chain traceability and food authenticity, and identify chemometrics combined with IoT as the principal enabling combination. Hassoun et al. [45] review Industry 4.0 technologies in the food industry and demonstrate that smartphone-based smart sensors can provide accurate, portable, and low-cost quality assessment. Both works are relevant here because traceability downstream of the farm gate depends on the provenance of the sensor data recorded upstream of it.

2.9. Security, Privacy, Governance, and Adoption

Zanella et al. [6] carry out an extensive study of security issues in smart agriculture and apply cybersecurity frameworks to IoT-specific weaknesses such as malware and the hijacking of autonomous systems. Gupta et al. [7] focus on security and privacy in smart farming and propose deep neural network-based malware detection to protect remote sensor data. Turning to the socio-economic dimension, Jakku et al. [46] examine the effects of digital farming and find that trust, transparency, and explicit data sharing agreements are prerequisites for widespread adoption by farmers. Ahmed et al. [47] summarise the current state of smart agriculture and conclude that public datasets and advanced sensors exist, but that comprehensive policy frameworks are still needed for broader adoption. Vijayakumar et al. [48] trace the development of agriculture from the first to the fourth revolution and contend that robotics and big data are necessary to address ecological problems such as declining soil health. Sharma et al. [49] apply Latent Dirichlet Allocation to the smart farming literature and identify the industrial revolution, climate-resilient farming, and the socio-technical effects of IoT adoption as the dominant research clusters.

2.10. Synthesis of the Reviewed Literature

Across these clusters, three points of consensus emerge. First, sensing, connectivity, and edge/cloud analytics are consistently treated as complementary rather than competing layers, with the choice of connectivity technology (LPWAN, cellular IoT, 5G, satellite) driven by the latency and coverage needs of the specific application [17,18,21]. Second, security and privacy are repeatedly identified as unresolved, cross-cutting concerns rather than solved engineering problems, spanning device-level authentication, malware resistance, and farmer data ownership [6,7,13]. Third, AI/ML methods are widely proposed for prediction and detection tasks, but the reviewed studies rarely report deployment-scale validation, explainability, or the labeled-data requirements needed for adoption [27,33,34]. Where the literature disagrees, it is mainly on architectural emphasis: some studies favor cloud-centric analytics for scale [20], while others argue that rural connectivity constraints make edge-first or offline-capable designs necessary [16,21]. These points of consensus and disagreement directly motivate the ten open problems identified in Section 4: they mark the boundary between what the field has converged on (the value of layered CPS architectures and the centrality of security and interoperability) and what remains unresolved (explainability, governance, labeled-data scarcity, and cost-driven adoption barriers).

3. Background and Conceptual Foundations

Smart agriculture systems integrate sensing technologies, communication networks, and computational platforms to support data-driven farm management [16,25]. These systems are commonly structured as cyber–physical architectures composed of multiple interacting layers. The layers comprise the physical processes, the sensing infrastructure, the communication networks, and the analytical platforms [11,17]. This section presents the conceptual building blocks of smart agriculture. It begins with system-level definitions, then covers the enabling technologies and reference architectures, the trade-offs between sensor data transmission networks, and finally the lifecycle that sensor data follows from collection to use.

3.1. Smart Agriculture in the 4IR as a Cyber–Physical System

Smart agriculture, often framed as "Agriculture 4.0", is best understood as an agricultural cyber–physical system (CPS). A CPS of this kind couples four elements in a feedback loop: the physical farm processes of soil, plant, animal, and machine; sensing and actuation; communications; and computation and AI [1,2]. The purpose of the coupling is to deliver time-sensitive functions under real-world uncertainty, which in agriculture arises from weather, biology, terrain, connectivity, and labour availability. This framing matches formal definitions of CPS, which emphasise the integration of computation, communication, sensing, and actuation with physical systems, together with time sensitivity and interaction with the environment and with humans [11,25].

Within the 4IR, CPS is widely recognised as a foundational construct. The 4IR is commonly characterised by the fusion of the physical and the digital through connected, intelligent systems, and CPS is one of its signature building blocks [1,45]. In agriculture, CPS manifests as layered architectures that span field devices, networks, edge and cloud platforms, analytics and AI, decisions, and actuators or robots. These layers are increasingly organised around interoperable data models [11,25]. Operationally, the CPS view clarifies three things. It explains why smart agriculture can improve yields and input efficiency, through better observability, targeted interventions, and automation. It explains how such systems fail, through sensor drift, model error, connectivity outages, cyberattacks, and unsafe autonomy. Finally, it identifies the engineering requirements that dominate, namely latency, uptime, resilience, and safety [6,32,37]. The view also makes the connectivity trade-off explicit. Low-Power Wide-Area Network (LPWAN) sensing, such as Long Range Wide Area Network (LoRaWAN) or Narrowband IoT (NB-IoT), is well suited to coverage and battery life but not to tight control loops. 5G-class connectivity can support low-latency robotics and high-rate video, but at a higher cost and with greater deployment complexity [18,19,21].

Security, privacy, and safety are cross-cutting CPS concerns. Agricultural CPS inherits operational-technology (OT) realities—performance and safety constraints, remote access pressures, and potentially severe "physical world" consequences (production losses, safety hazards). Mitigations therefore combine IoT baselines, OT segmentation, secure-by-design components, and sector safety standards for automation [6,7]. Modern smart agriculture systems therefore rely on layered architectures where field-level sensing devices communicate with gateways and platforms, enabling analytics, machine learning models, and decision-support tools [11,17]. These architectures support applications ranging from irrigation scheduling to pest detection, crop monitoring, and autonomous machinery coordination [16,38].

3.2. Role of Sensor Data in Field-Based IoT Environments

Agricultural data are primarily generated through distributed sensing systems deployed directly in the field. Unlike many industrial environments, agricultural sensing infrastructures must operate in remote and often harsh environmental conditions for extended periods with limited maintenance [13,16].

The CPS concept is central in 4IR narratives capturing the defining shift from "digital tools supporting operations" to "digital intelligence embedded into operations." In this framing, CPS are the mechanisms by which sensing, networking, and computation become continuously entangled with physical processes, enabling new capabilities such as automation, remote operation, optimization, and new business models [1].

CPS integrate computation, communication, sensing, and actuation with physical systems in order to fulfil time-sensitive functions. The degree of environmental and human interaction varies between deployments. This model is particularly suitable for agriculture for two reasons. Farms are open systems subject to strong exogenous drivers such as weather, pests, and market constraints. In addition, many agricultural workflows, including irrigation, disease response, and harvest logistics, are both time-sensitive and seasonally constrained [16,35]. In agriculture-specific literature, Agriculture 4.0 explicitly frames farming as a CPS cycle (observe → decide → act) across heterogeneous technologies and stakeholders, often using terms such as agricultural cyber–physical systems (ACPS) [25,36].

Sensor data streams typically include environmental variables such as soil moisture, temperature, humidity, crop condition indicators, and machine telemetry [12,20]. These measurements form the basis for operational decision-making, making the reliability and interpretation of sensor data critical to the performance of agricultural cyber–physical systems. As these systems operate as closed feedback loops, failures in sensing, communication, or processing stages can directly influence physical outcomes such as irrigation levels, fertilization rates, or crop protection measures [8,9]. A useful engineering abstraction is a layered CPS model. Such a model preserves the closed-loop nature of farming operations while making the interfaces between layers explicit [11,25]. The mapping in Table 1 combines the core elements of the CPS definition, namely sensing, communication, computation, and actuation, with the practical reality that farms operate as systems of systems composed of machines, fields, barns, and supply-chain interfaces [17,40].

Table 1 depicts the analysis of the CPS layers and how they interact, alongside their components and the corresponding constraints [11,25]. Figure 1 illustrates the fundamental cyber–physical feedback loop that underpins smart agriculture systems. Measurements obtained from sensors are transmitted through the communication and computation layers, where decision logic processes the observed state and generates control actions. These actions are then applied to the physical environment through actuators, closing the loop and enabling continuous adaptation to changing agricultural conditions.

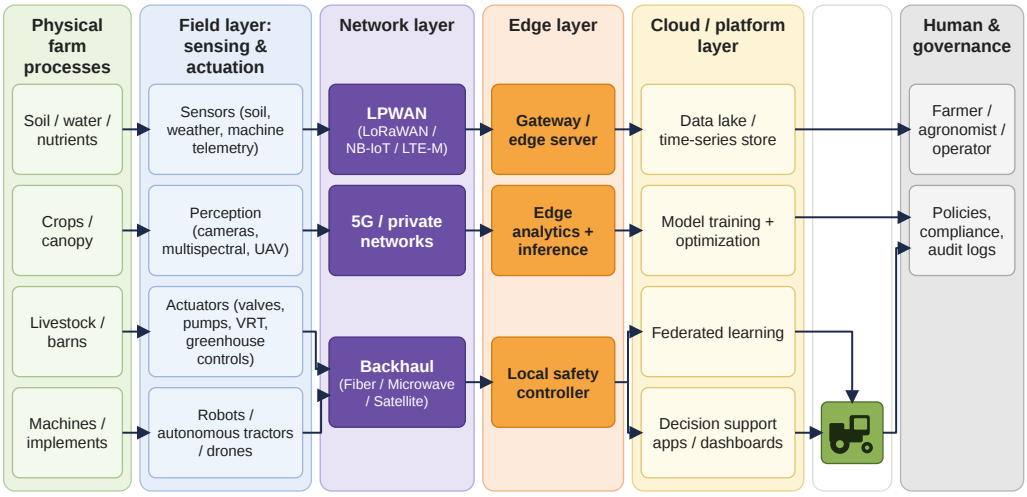


Figure 1. Reference data flow diagram of the agricultural CPS closed loop. The diagram shows the information flow from the physical farm processes through sensing, network, edge, and cloud layers, and the corresponding action flow back to the field, with both coordinated by the decision logic.

Table 1. CPS layers and how they interact in the feedback loop [1,11,25].

CPS Layer	Role in Feedback Loop	Smart Agriculture Components	Engineering Constraints
Physical processes	Real-world processes being controlled	Soil moisture, crops, livestock physiology, microclimate, machinery dynamics	Uncertainty, seasonality, biological nonlinearity
Sensing and perception	Observability of system state	Sensors, weather stations, cameras, Unmanned Aerial Vehicle (UAV) imagery, telemetry	Calibration drift, missing data, power limits
Actuation and mobility	Changes physical state	Irrigation valves, variable-rate applicators, robots, autonomous tractors	Safety, timing, maintainability
Communications	Transfers data and commands	Gateways, LoRaWAN, NB-IoT, Long-Term Evolution for Machines (LTE-M), 5G, WiFi, satellite	Coverage, latency, cost
Edge/fog compute	Local processing and control	On-machine controllers, edge servers, Multi-access Edge Computing nodes	Low latency, connectivity tolerance
Cloud/platform compute	Global optimization and analytics	Data lakes, Machine Learning (ML) training, fleet analytics	Bandwidth, data governance
Applications and integration	Presentation and system integration	Dashboards, farm management information systems, traceability systems	Data model validity, interoperability
Human/organizational control	Oversight and policy	Farmers, agronomists, regulators	Trust, explainability

3.3. Enabling Technologies and Reference Architectures

A typical smart agriculture architecture includes field-level sensing devices that collect environmental measurements, gateways that aggregate and transmit data, and cloud or edge computing platforms that process and analyze the information [16,17,25]. The resulting insights support precision interventions such as variable-rate irrigation, fertilization, and crop protection strategies [37,39]. Ensuring that sensor data remains trustworthy across this pipeline is essential for achieving the operational benefits associated with precision agriculture [8,9]. Connectivity, compute placement, and data/AI pipelines dominate most agricultural CPS designs:

- IoT sensing and device management underpin observability [13,15]. They also introduce device security and lifecycle management requirements, and manufacturer-focused IoT security baselines are increasingly treated as procurement expectations in CPS supply chains [6,7].
- 5G-era requirements, covering latency, reliability, positioning, and flexible deployment through slicing and edge hosting, are motivated by vertical use cases such as UAV control and automation [4,21]. Agriculture, however, typically mixes these with lower-cost LPWAN telemetry [18,19].
- Edge computing is repeatedly emphasized in agriculture because rural connectivity is variable and because time-sensitive tasks (local control, video analytics) benefit from nearby processing with reduced bandwidth use [10,22,23].
- Blockchain appears most credibly in traceability and supply-chain integrity contexts rather than as a universal farm database [44,45]. It can address multi-stakeholder trust issues, but faces adoption barriers relating to standardization, governance, and incentives [46,47].

- Federated learning is increasingly proposed to reduce privacy risks and data centralization while enabling cross-farm model improvement [7,34]. It nevertheless introduces communication overhead, data heterogeneity issues, and new attack surfaces [6,33].

The technologies discussed above are often deployed within broader architectural patterns that determine where data are processed, where decisions are made, and how information flows between field devices, edge infrastructure, and cloud services [10,25]. Each pattern offers different trade-offs in terms of latency, bandwidth consumption, operational resilience, deployment complexity, and scalability. Table 2 compares the four principal architectural approaches reported in the smart-agriculture literature, namely the cloud-centric, edge-first, hybrid edge–cloud, and federated patterns. For each of them, the table contrasts the data flow and locus of control, the combined latency and bandwidth profile, the resilience profile, the cost and complexity of deployment, and the agricultural scenarios for which the pattern is best suited. As shown in Table 2, no single architectural pattern is universally optimal. Cloud-centric approaches provide strong analytical capabilities but depend heavily on network connectivity [20], whereas edge-first designs improve resilience and support time-critical control decisions in rural environments with intermittent connectivity [10,22]. Hybrid edge–cloud architectures attempt to balance these competing requirements [21], while federated approaches add capabilities related to coordination, privacy preservation, and cross-farm collaboration [7,46]. These observations motivate the trust-oriented layered architecture proposed later in this paper, which combines the strengths of multiple architectural patterns while addressing their limitations through integrated security, interoperability, provenance, and governance mechanisms.

Table 2. Architectural patterns in smart-agriculture cyber–physical systems [7,10,21,25].

Architecture Pattern	Data Flow & Control Locus	Latency & Bandwidth Profile	Resilience Profile	Cost/Complexity Profile	Best-Fit Scenarios
Cloud-centric [20,25]	Raw/aggregated data → cloud; decisions back to field	Often higher, depending on Wide Area Network (WAN) availability; higher uplink demand	Vulnerable to WAN outages	Lower on-site complexity; higher cloud dependency	Advisory analytics, reporting, historical optimization
Edge-first (local control) [10,22,23]	Sense → edge inference/control; cloud for sync	Low for local loops; uplink reduced by local filtering	Stronger under intermittent WAN	Higher on-site ops needs	Irrigation control, safety interlocks, real-time alarms
Hybrid edge–cloud [19,21]	Split: edge executes; cloud trains/optimizes	Balanced, with critical paths local; moderate uplink	Better than cloud-only; still needs sync	Moderate-to-high integration effort	Video + AI (pest/weed), fleet robotics + monitoring
Federated/data-space oriented [7,46]	Local data stays local; shared models/metadata	Depends on coordination granularity; can reduce raw data sharing	Governance-dependent; reduces single-point cloud risk	High governance and interoperability demands	Cross-farm benchmarking, privacy-sensitive ML, cooperative platforms

3.4. Transmission Networks and Connectivity Trade-Offs

Connectivity choices strongly shape what can be controlled in a closed loop and what can only be monitored [17,18]. The conceptualization below is drawn from publicly available standards documentation of the 3rd Generation Partnership Project (3GPP) and the International Mobile Telecommunications-2020 (IMT-2020) requirements, rather than from peer-reviewed agricultural studies. They should be read as indicative design targets and not as field-measured performance. Where the agricultural literature reports comparable behaviour, it is cited alongside [18,19,21]:

- IMT-2020 minimum performance requirements explicitly distinguish ultra-low user-plane latency targets for Ultra-Reliable Low-Latency Communications (1 ms) versus

Enhanced Mobile Broadband (4 ms), and define dense device connectivity targets for Massive Machine-Type Communications-class scenarios.

- Cellular IoT designs (enhanced Machine-Type Communication/LTE-M and NB-IoT) were developed to emphasize long battery life (on the order of 10 years under appropriate traffic/coverage assumptions) and deep coverage (expressed via maximum coupling loss targets); comparable battery-life gains are reported for LPWAN deployments in agriculture [18,19].
- Satellite integration is increasingly formalized in 5G-era standards evolution (e.g., Non-Terrestrial Networks (NTN) as a major Release 17 focus), but latency and service characteristics depend strongly on orbit and architecture (e.g., Geostationary Earth Orbit vs Low Earth Orbit), a dependency that is also noted in agricultural connectivity reviews [21,28].

These trade-offs are therefore relative rather than absolute, and they follow from the capability emphasis documented in the IMT-2020 requirements, the 3GPP cellular IoT summaries, and the NTN discussions. The practical consequence is architectural. Farms often blend LPWAN-class telemetry, which offers low power consumption and wide coverage, with higher-capacity links such as private 5G, Wi-Fi, or satellite backhaul for robotics, imaging, and time-sensitive control. A single network rarely fits all agricultural traffic classes [17,20,21].

WSNs form the sensing backbone of many smart agriculture deployments [16,17]. These networks consist of distributed sensor nodes that monitor environmental variables and communicate measurements to gateways or base stations. Agricultural deployments typically cover large geographic areas and must operate for long periods. This introduces challenges related to power consumption, connectivity reliability, and physical accessibility [13,16]. Sensor nodes may experience environmental interference, hardware degradation, or communication disruptions that affect data availability and quality [9].

3.5. Sensor Data Lifecycle in Smart Agriculture

Sensor data acquisition occurs directly at the interface between digital systems and physical agricultural environments [13,15]. Measurements must accurately capture environmental conditions while accounting for hardware limitations and environmental variability. vCalibration drift and environmental noise degrade measurement reliability over time [9]. Missing contextual information, such as location or timestamp, reduces the interpretability of the collected data [8,11]. Maintaining measurement quality therefore requires careful management of sensor calibration, metadata, and environmental context.

Once collected, sensor data must be transmitted through communication networks to aggregation points or computing platforms [17,20]. Agricultural networks often employ a combination of low-power wide-area communication technologies and gateway-based aggregation to transmit data efficiently [18,19]. Edge computing can be used to filter or preprocess measurements locally before forwarding them to cloud-based platforms for long-term storage and analytics [10,22].

Collected data are typically stored in centralized repositories where they can be processed, analyzed, and reused for multiple purposes [11,41]. Historical datasets enable trend analysis, predictive modeling, and optimization of agricultural operations [32,33]. However, these applications depend on the availability of reliable and well-documented data, which is why traceability and auditability matter across the whole data lifecycle [8,44].

4. Open Problems in Smart Agriculture Sensor Data Ecosystems

The open problems and the corresponding conceptual solutions discussed in this work were derived through the narrative literature synthesis described in Section 2.1, clustering

recurring themes and explicitly stated limitations across the reviewed works, combined with design-oriented reasoning about how established techniques (semantic middleware, Physical Unclonable Function (PUF) based identity, Conflict-Free Replicated Data Type (CRDT) based synchronization, explainable AI, federated learning, self-sovereign identity) could plausibly be composed to address each problem. The approach follows neither a systematic review protocol nor an empirical evaluation. It positions a conceptual framework that rests on the literature synthesis reported above, and that scope limitation is revisited in Section 7. The proposed approach consolidates recurring challenges reported in the literature, formulates ten open problems, and integrates corresponding solution patterns into a unified trust-oriented reference architecture for future smart-agriculture systems.

4.1. Derivation of the Open Problems

The open problems carry the remainder of the paper: each is answered by a conceptual solution in Section 5 and is assigned to a layer of the architecture in Section 6. The procedure by which they were formulated is therefore stated explicitly here, rather than left implicit in the narrative synthesis.

The procedure has four steps:

1. **Extraction:** Each of the works retained at full-text stage under the protocol of Section 2.1 was read for statements of unresolved difficulty. Three kinds of statement were extracted: a limitation the authors named in their own work, a gap or open challenge they named in the field, and a direction they identified as required future research. Statements of achievement were not extracted. Each extracted statement was recorded together with the work it came from and the thematic cluster of Section 2 in which that work is discussed.
2. **Coding:** The extracted statements were coded inductively, in the sense that the category labels were taken from the statements themselves rather than from a pre-existing taxonomy. Statements expressing the same difficulty in different vocabulary were merged into one category. Statements confined to a single crop, a single site, or a single vendor implementation were treated as instances of a category rather than as categories in their own right.
3. **Retention:** A category was retained as an open problem only if it satisfied three conditions. Condition (i): it is stated as unresolved in at least three of the reviewed works, drawn from more than one thematic cluster, so that it does not rest on a single line of work. Condition (ii): it persists at the level of the system rather than the component, in the sense that it cannot be removed by substituting one device, protocol, or model for another. Condition (iii): it can be located at a definite layer of the CPS model of Table 1, so that any solution to it has a well-defined place in an architecture. Categories that failed condition (i) were kept as observations in Section 2.10 but not promoted to problems; categories that failed condition (ii) were folded into the category of which they were an instance; categories that failed condition (iii) were predominantly agronomic rather than informational and fall outside the scope of this paper.
4. **Ordering:** The retained categories were ordered along the path taken by the data through the system, from the device layer to the socio-economic layer, as set out in Section 4.2.

Ten categories satisfied all three retention conditions. Table 3 records, for each of them, the thematic cluster from which it was drawn, the reviewed works whose stated limitations or future-work statements support it, and the CPS layer at which it arises. The table is the audit trail for the list that follows: every problem stated in Section 4.2 can be traced through

it back to the works that motivate it, and the citations attached to the problem statements themselves are the same works.

Three qualifications attach to this procedure. First, it inherits the limitations of the search protocol of Section 2.1. The underlying corpus is a judgement sample rather than an exhaustive one, so the ten problems are those that are salient in that corpus, and not the complete set of open problems in the field. Second, extraction and coding were carried out by the author team without independent duplication, so no inter-rater agreement can be reported and the boundaries between categories reflect the authors' reading. Third, the threshold of three independent works in condition (i) is a convention adopted for this synthesis rather than a validated criterion. A different threshold would yield a somewhat different list, most plausibly by admitting further problems rather than by removing any of these ten, each of which is supported by more works than the threshold requires.

Table 3. Evidence base for the ten open problems. For each problem, the table gives the thematic cluster of Section 2 from which it was drawn, the reviewed works that state the corresponding difficulty as a limitation or as required future work, and the layer of the CPS model of Table 1 at which it arises.

#	Open Problem	Literature Cluster	Supporting Works	CPS Layer (Table 1)
1	Physical security and durability of edge sensors	Sensing hardware (Section 2.2); security (Section 2.9)	Aarif et al. [13]; Ayaz et al. [16]; Mansoor et al. [23]; Zou et al. [9]; Zanella et al. [6]	Sensing and perception
2	Lightweight security and authentication for constrained devices	Security and privacy (Section 2.9)	Gupta et al. [7]; Zanella et al. [6]; Aarif et al. [13]	Sensing and perception; communications
3	Internet connectivity and infrastructure in rural areas	Connectivity (Section 2.3); edge (Section 2.4)	Aldhaheri et al. [19]; Polymeni et al. [21]; Shi et al. [10]; Karunathilake et al. [37]; Vinod Chandra et al. [36]	Communications; edge/fog compute
4	Technical limitations of UAVs	Remote sensing and UAV platforms (Section 2.5)	Phang et al. [27]; Guebsi et al. [29]; Delavarpour et al. [30]; Maddikunta et al. [4]; Khanal et al. [28]	Sensing and perception; actuation and mobility
5	Data interoperability and standardization	Edge and IoT architectures (Section 2.4); connectivity (Section 2.3)	Aldhaheri et al. [19]; Qazi et al. [22]; Mowla et al. [17]; Araújo et al. [25]; Jakku et al. [46]	Applications and integration
6	Scarcity of labeled training data for machine learning	Data analytics and AI (Section 2.6)	Phang et al. [27]; Linaza et al. [34]; Garg and Alam [33]; Vijayakumar et al. [48]	Cloud/platform compute
7	Explainability and trust in AI models	Data analytics and AI (Section 2.6); governance (Section 2.9)	Linaza et al. [34]; Garg and Alam [33]; Mansoor et al. [23]; Jakku et al. [46]	Human/organizational control
8	Cross-cloud trust and federated access control	Security, privacy, and governance (Section 2.9)	Gupta et al. [7]; Jakku et al. [46]; Ahmed et al. [47]	Cloud/platform compute; human/organizational control
9	Data ownership, privacy, and ethical concerns	Security, privacy, and governance (Section 2.9)	Jakku et al. [46]; Ahmed et al. [47]; Vijayakumar et al. [48]; Aarif et al. [13]	Human/organizational control
10	High initial costs and the digital divide	Field applications (Section 2.7); governance (Section 2.9)	Karunathilake et al. [37]; Sharma and Shivandu [38]; Vijayakumar et al. [48]; Ahmed et al. [47]	Human/organizational control

4.2. Open Problems

The ten problems below are ordered so that they follow the path of the data through the system. Problems 1 and 2 concern the sensing device itself, in terms of physical survival and of cryptographic identity. Problems 3 and 4 concern the transmission network and the mobile sensing platforms that depend on it. Problems 5 to 7 concern the data and the intelligence derived from it, moving from representation to training data and then to the

interpretability of the resulting models. Problems 8 to 10 concern governance, ownership, and the socio-economic conditions for adoption. Section 5 addresses the ten problems in exactly this order, and Table 4 maps each problem to its conceptual solution and to the architectural layer at which that solution operates.

1. **Physical Security and Durability of Edge Sensors:** Unlike sensors in smart cities, agricultural IoT devices are deployed in open fields where they are continuously exposed to harsh environmental conditions, wildlife, and heavy farming machinery [6]. Ensuring the physical integrity, preventing sensor degradation from dirt and corrosion, and maintaining calibration over time without utilizing expensive tamper-resistant protective boxes remain significant practical challenges [13,23].
2. **Lightweight Security and Authentication for Constrained Devices:** Traditional cryptographic schemes and public-key infrastructures are too computationally intensive for the resource-constrained, low-power IoT sensors typically used in agriculture. Developing scalable, lightweight, and energy-efficient authentication and encryption methods requiring minimal human intervention remains a critical research area [6,7].
3. **Internet Connectivity and Infrastructure in Rural Areas:** Deploying smart agriculture technologies is severely hindered by the lack of reliable internet connectivity and high-bandwidth networks in remote and rural farming regions [36,37]. To address high network latency and limited bandwidth, future research must focus on developing advanced edge computing architectures that allow for localized, real-time data processing without relying exclusively on centralized cloud infrastructure [10,19,21].
4. **Technical Limitations of UAVs:** Despite their benefits for aerial monitoring, the widespread use of drones is restricted by significant technological limitations, including short battery life, limited payload capacity, and susceptibility to adverse weather conditions [27,29]. Furthermore, navigating complex regulatory constraints and achieving automated coordination for "drone swarms" represent major ongoing challenges [27,29,30].
5. **Data Interoperability and Standardization:** One of the main concerns is the significant lack of unified data formats, standard communication protocols, and open Application Programming Interfaces (APIs) across the heterogeneous devices and platforms used in smart farming [19,22]. This heterogeneity complicates the seamless integration of agricultural data, making it difficult to scale systems and establish interoperable architectures across different vendors [17,19,46].
6. **Scarcity of Labeled Training Data for Machine Learning:** The successful application of supervised machine learning in agriculture is currently bottlenecked by the lack of large, accurately labeled datasets [27]. High-quality training data are required to build robust models, especially for detecting rare events or specific plant diseases, and the integration of domain-specific semantic models is needed to improve AI precision when training data is small [34,48].
7. **Explainability and Trust in AI Models:** While AI and deep learning models offer precise performance for agricultural predictions, they often act as "black boxes", making it difficult for farmers to understand how specific decisions are made [33,34]. The lack of explainability reduces farmers' trust in AI technologies, creating a need for interpretable AI models that clearly communicate data usage and algorithmic decision-making [23,34].
8. **Cross-Cloud Trust and Federated Access Control:** When farmers borrow smart equipment from one another or use analytics services from multiple third-party cloud providers, establishing trust and access control across federated environments becomes highly complex. Determining how to automatically delegate, revoke, and securely

manage access rights for temporary laborers and borrowed physical or virtual assets is an open problem requiring advanced, farm-specific access control models [7].

9. **Data Ownership, Privacy, and Ethical Concerns:** The massive generation of agricultural data raises severe concerns over data privacy, security, and ownership [47]. Farmers are often reluctant to share their data due to unclear legal frameworks, complex contracts, and the fear that large corporations or third-party providers might exploit their data for price discrimination or sell it without fair compensation [13,46,48].
10. **High Initial Costs and the Digital Divide:** The substantial capital investment required to purchase and maintain smart sensors, automated machinery, and data analytics tools creates a major barrier to adoption, particularly for small-scale and resource-limited farmers [37,38]. This financial hurdle, coupled with a lack of digital literacy and technical training, threatens to widen the technological gap between large industrialized farms and smallholder farmers in developing nations [37,48].

Table 4. Mapping of the ten open problems of Section 4 to the conceptual solutions of Section 5 and to the architecture layers of Section 6.

#	Open Problem (Section 4)	Conceptual Solution (Section 5)	Architecture Layer
1	Physical security and durability of edge sensors	Self-healing modular sensor nodes	Sensing and Device (L_1)
2	Lightweight security and authentication for constrained devices	Hierarchical edge security	Sensing and Device (L_1)/Connectivity and Edge (L_2)
3	Internet connectivity and infrastructure in rural areas	Offline-first edge/fog architecture	Connectivity and Edge (L_2)
4	Technical limitations of UAVs	Hybrid UAV fleet management framework	Sensing and Device (L_1)
5	Data interoperability and standardization	Semantic agricultural middleware layer	Interoperability and Data Integration (L_3)
6	Scarcity of labeled training data for machine learning	Federated agro-learning network	Intelligence and Decision Support (L_4)
7	Explainability and trust in AI models	Explainable decision support dashboard	Intelligence and Decision Support (L_4)
8	Cross-cloud trust and federated access control	Decentralised identity federation	Governance, Ownership, and Trust (L_5)
9	Data ownership, privacy, and ethical concerns	Farmer-centric data sovereignty platform, with simplified permission templates and legal validation	Governance, Ownership, and Trust (L_5)
10	High initial costs and the digital divide	Smart Agriculture as a Service	Governance, Ownership, and Trust (L_5)

5. Conceptual Solutions to Open Problems

This section presents targeted, framework-level solutions to each of the ten identified open problems, followed by a unified architectural synthesis (Section 6) that integrates these solutions into a coherent, end-to-end system. The subsections appear in the same order as the open problems of Section 4.2, so that Solution n addresses Problem n. The correspondence is intended to be one-to-one and exhaustive in both directions: every problem in Section 4.2 is answered by exactly one subsection below, and no subsection below answers a problem that is not in that list. Two consequences of this discipline should be flagged, because they are the points at which the two lists could otherwise appear to

diverge. First, a solution may operate at more than one architectural layer even though it answers a single problem: Solution 2 (see, Section 5.2) is the case in point, since its mechanisms are split between the node and the gateway, and Table 4 records both layers in that row. Second, the discussion of farmer usability and legal validation, which in the previous version stood as a separate item without a corresponding problem, is a subsubsection of Solution 9 (see, Section 5.9), because it concerns the same problem of data ownership and consent rather than a further one. The solutions below are conceptual design proposals synthesizing established techniques and standards (e.g., PUF-based device identity, Advanced Encryption Standard (AES)/Hash-based Message Authentication Code (HMAC) primitives, CRDT-based synchronization, SHapley Additive exPlanations (SHAP) based explainability, federated averaging, World Wide Web Consortium Decentralized Identifiers) into an integrated architecture for smart-agriculture data pipelines. The proposed components have not been implemented or empirically evaluated in this work. The proposal should therefore be read as a design blueprint and as a set of testable hypotheses for future work [6,7,34].

5.1. Solution 1: Self-Healing Modular Sensor Nodes for Harsh Field Environments

Agricultural IoT devices are routinely exposed to soil corrosion, water ingress, mechanical shock from farming machinery, and sustained ultraviolet radiation. These degradation pathways substantially shorten sensor operational lifetimes and compromise measurement accuracy [6,9,13,16]. For this purpose, a Self-Healing Modular Sensor Node design that addresses both physical resilience and long-term calibration integrity could be designed. The hardware enclosure conforms to Ingress Protection 67 standards, providing complete dust exclusion and resistance to temporary water submersion. The functional sensing element could be implemented as a hot-swappable cartridge that can be replaced by non-specialist field personnel without tools, minimising downtime and repair cost. Energy is supplied through a combination of solar trickle charging and a low-self-discharge primary cell, enabling uninterrupted operation throughout overcast periods. Additionally, a predictive maintenance subsystem continuously monitors battery terminal voltage, internal temperature, and measurement drift indicators.

When sensor readings diverge from a statistical baseline, the node initiates an automatic multi-point recalibration routine using stored reference values and, if necessary, triggers an alert to the farm management dashboard. Finally, further accelerometer data could be analysed to detect vibration signatures consistent with tampering or accidental displacement, providing physical security logging without additional hardware. These features could reduce maintenance visits, lower lifetime ownership cost, and sustain data quality across multi-year deployments [13,23]. These are design expectations rather than findings. None of them has been measured here, and Section 6.4 states the accelerated-ageing and multi-season protocol by which they would be tested.

5.2. Solution 2: Hierarchical Edge Security for Constrained IoT Devices

Conventional public-key cryptographic schemes impose computational and energy costs that exceed the capabilities of the low-power microcontrollers commonly embedded in agricultural sensors [6,7,13]. A hierarchical edge security model could be adopted, distributing security responsibilities across three tiers, each sized to its available resources. The threat model considered in this study follows the Dolev–Yao adversary model [50] and includes the following attack scenarios:

- **Active man-in-the-middle (MITM) attacks at gateway:** The adversary is assumed to be capable of intercepting, modifying, or injecting messages between sensor nodes and the edge gateway.

- **Node physical capture:** The adversary may physically capture or steal a sensor node and gain access to its stored credentials and local data for a period of up to 30 days.
- **Replay attacks:** The adversary may record legitimate messages and attempt to retransmit them within a multi-hour time window (e.g., 1–6 h, an illustrative value that would need to be tuned to the target deployment) in order to trigger unauthorized or outdated system actions.
- **Sybil attacks:** The adversary may create multiple false node identities to inject misleading measurements or disrupt trust relationships within the sensor network.
- **Cloud service compromise:** The adversary may compromise the cloud service and exfiltrate stored data, but is not assumed to be capable of modifying the deployed algorithms or decision-making logic.

The design targets four security guarantees, namely confidentiality, authentication, resistance to replay, and message integrity. Confidentiality is achieved through Advanced Encryption Standard with a 128-bit key in Counter Mode (AES-128-CTR) encryption, which protects transmitted data against eavesdropping. Counter Mode is a stream-cipher construction and therefore carries two conditions that the design must satisfy explicitly. First, it provides confidentiality only: the ciphertext is malleable, and a bit flipped in transit flips the corresponding plaintext bit undetectably unless a separate integrity mechanism is applied. The design therefore composes AES-128-CTR with HMAC in encrypt-then-MAC order, so that the tag is computed over the ciphertext and the associated metadata, and the gateway verifies the tag before attempting decryption. Second, the security of Counter Mode fails completely if a counter block repeats under the same key. The counter block is therefore constructed from the node identifier, a session-key epoch counter held in non-volatile memory, and the per-message sequence number, so that uniqueness survives node reboots and battery replacement, which are frequent events in field deployments [51]. Although confidentiality may be affected if a sensor node is physically captured, the cryptographic key is stored using a PUF, making key extraction significantly more difficult.

Authentication is provided through X.509 certificates at the gateway, which prevent impersonation attacks, while PUF-based node identities improve resistance against device cloning. Nevertheless, known PUF side-channel vulnerabilities must be considered and mitigated. Two further properties of PUF-based identity should be stated precisely rather than assumed. First, a PUF response is noisy: it varies with temperature, supply voltage, and device ageing, so it cannot be used directly as a key. A fuzzy extractor with an error-correcting code is required to derive a stable key from the raw response, and the associated helper data, although it need not be secret, must be integrity-protected, because an adversary able to modify helper data can bias the reconstructed key.

Second, resistance to Sybil attacks does not follow from the PUF alone. It follows from enrollment: each node is registered at commissioning time with a set of challenge–response pairs held by the gateway, and an identity that is not in that register cannot be authenticated. The adversary is thereby prevented from minting new identities, but is not prevented from continuing to use the identity of a node it has physically captured, which is why capture is handled by revocation rather than by the identity mechanism. Resistance to replay rests on two mechanisms that operate on different quantities and should not be conflated. The timestamp defines a freshness window: a message whose timestamp lies further from the gateway clock than the tolerance Δ is treated as stale and is discarded before it reaches the edge decision logic. This test bounds how long a captured message remains useful to an adversary, but it does not by itself detect a replay inside the window, and it presupposes loose time synchronisation between node and gateway, which the design obtains from a periodic time beacon issued by the gateway and authenticated under the same session key.

Detection inside the window is the task of the second mechanism: the gateway maintains, for each enrolled node, the highest sequence number it has accepted together with a sliding bitmap of recently seen values below that maximum, and rejects any message whose sequence number is either already recorded or below the window floor. The sequence number is monotonic and is persisted in non-volatile memory on the node, so that it does not restart after a reboot. The window tolerance Δ is thus an availability parameter, traded against clock drift on the node, whereas the sequence-number state is what supplies the replay guarantee.

Integrity protection is provided by HMAC with the Secure Hash Algorithm 256-bit (HMAC-SHA-256). Two points of precision are needed here, because this guarantee is easily overstated. First, the security of HMAC rests on a pseudorandom-function assumption on the keyed compression function rather than on collision resistance of the bare hash [52], so the message-authentication property is retained under assumptions weaker than those required for collision resistance.

Second, the probability that a blind forgery is accepted is approximately 2^{-t}, where t is the length in bits of the tag actually transmitted, and is not 2^{-256} unconditionally. Constrained LPWAN payload budgets frequently force the tag to be truncated, and a deployment that truncates to $t = 64$ bits accepts a bound of 2^{-64} per attempt. The tag length is therefore a deployment parameter that must be stated rather than assumed, and the present design fixes a minimum of $t = 128$ bits for agricultural telemetry. Neither bound applies to an adversary in possession of the key, which is the node-capture case, and both are generic properties of the construction rather than values measured within this work.

Despite its advantages, the proposed security design has several limitations. First, compromise of the edge gateway may expose downstream data because the architecture does not provide full end-to-end encryption between sensor nodes and the cloud. This design choice represents a deliberate trade-off, as edge processing requires access to decrypted measurements for filtering, aggregation, anomaly detection, and real-time decision support. A second limitation concerns node capture. The PUF protects the key at rest, in the sense that it is not stored in readable non-volatile memory, but it does not protect a running device: an adversary with physical possession can exercise the fuzzy extractor and obtain the derived key. Within the assumed capture period of up to 30 days, the adversary can therefore produce messages that are valid for that node until the node is revoked, and the effective bound on the damage is the detection-and-revocation delay rather than any cryptographic quantity. The design consequently requires the session-key rotation epoch and the gateway's node-liveness audit to be shorter than the expected time to notice a missing node, and it places revocation, through an allow-list maintained at the gateway and mirrored in the cloud trust log, on the critical path of the security argument. Node capture also degrades PUF reliability over time, as a result of environmental stress, device aging, temperature variation, and repeated operation. Consequently, periodic recalibration may be necessary, introducing additional maintenance costs that are nevertheless considered acceptable when compared with the cost of replacing or upgrading deployed sensor nodes. Finally, the proposed mechanism does not fully protect against protocol-level attacks such as forged calibration commands or malicious configuration updates. This limitation is acceptable in the target agricultural deployment because such commands occur infrequently, generally require gateway-level authorization, and can be further controlled through operational procedures, manual validation, and audit logging.

Regarding the comparison with alternatives, several alternative security approaches were considered before selecting the proposed design. A full Public Key Infrastructure (PKI) offers strong authentication and scalable certificate management but incurs substantial

computational and energy overhead due to asymmetric cryptographic operations, making it less suitable for low-power, coin-cell-operated sensor nodes that require long-term autonomous operation. Symmetric key management schemes are computationally efficient and well suited to resource-constrained devices; however, they create significant challenges for key distribution and rotation, particularly in large-scale deployments and in threat models that include physical node capture. Hardware Security Modules provide robust hardware-based key protection and secure cryptographic execution, but their per-unit hardware cost is substantially higher than that of a PUF circuit integrated into an existing microcontroller, which is often prohibitive at the scale of large agricultural deployments; a precise cost comparison would require vendor-specific quotes and is left for future work. In contrast, the proposed PUF-based solution offers a lower-cost alternative while still enhancing resistance to key extraction and device cloning attacks.

The security mechanisms are finally analysed across the three tiers of node, gateway, and cloud. At the sensor node tier, symmetric encryption using AES-128 with periodically rotated session keys provides confidentiality with minimal processing overhead. Node identity is established through PUFs, which exploit manufacturing variations in silicon to generate hardware-intrinsic identifiers that are substantially harder to clone or extract than stored keys [6,7]. This removes the need for secure key-storage hardware, reducing unit cost. At the gateway tier, a more powerful microprocessor manages X.509 certificate lifecycle, performs mutual authentication with both sensors and the cloud endpoint, and enforces message integrity checks before forwarding. At the cloud tier, a trust-log service records authentication events and feeds an anomaly detection engine that flags irregular communication patterns indicative of node compromise or replay attacks.

The trust boundary of this decomposition should be stated precisely, because "end-to-end" means different things for the two properties involved and the distinction determines what the design does and does not deliver. Confidentiality is hop-by-hop, not end-to-end: the gateway terminates the node-to-gateway channel and re-encrypts towards the cloud, and it necessarily holds plaintext, because filtering, aggregation, anomaly detection, and local control all operate on decrypted measurements. The gateway is therefore inside the trust boundary by construction, and its compromise exposes the measurements passing through it, as recorded among the limitations above. Integrity and origin authentication, by contrast, can be end-to-end and are treated as such in this design: the tag computed on the node over the canonical measurement record is retained and forwarded with the record through the middleware to the cloud, so that a consumer of the data at the cloud tier can verify that a given measurement originated at the claimed node and was not altered at or after the gateway. Aggregates computed at the gateway cannot inherit this property and are instead signed by the gateway under its own certificate, so that the provenance record distinguishes a node-attested measurement from a gateway-attested derivative. This separation is what makes the provenance chain of Section 5 meaningful: the lineage identifier of Solution 5 records which of the two attestations a record carries. A design in which the cloud must also be excluded from the trust boundary would require end-to-end encryption from node to cloud and would forgo edge processing altogether, or would require homomorphic or trusted-execution techniques whose cost on coin-cell hardware is prohibitive at present. The decomposition described here therefore trades confidentiality at the gateway for local autonomy, while keeping integrity verifiable along the whole path, and it remains consistent with the energy and cost constraints of field-deployed agricultural hardware.

5.3. Solution 3: Offline-First Edge/Fog Architecture for Rural Connectivity

Since the unreliability of internet infrastructure in remote agricultural regions renders cloud-dependent architectures impractical for real-time field operations [10,21,36,37], an Offline-First Edge/Fog Architecture could be deployed, where every farm is equipped with a solar-powered edge gateway that operates as a fully autonomous processing node. When cloud connectivity is restored, the edge gateway synchronizes locally buffered data with the cloud repository. Updates may occur independently at the edge and in the cloud. A gateway may, for instance, record new sensor data while cloud-based model retraining is in progress. Synchronisation can therefore lead to data conflicts. To ensure deterministic and convergent reconciliation, the synchronisation layer is specified as a Conflict-Free Replicated Data Type [53]. Two distinctions have to be drawn before the mechanism is stated, because the term is sometimes applied loosely to any timestamp-based merge, and the convergence guarantee does not survive that usage.

The first distinction is between the two classes of data that the gateway holds, which require different replicated types. Sensor oservations are immutable facts, each identified by the triple (node identifier, measured quantity, observation time). Two replicas can never hold different values for the same triple, because a measurement is not revised once taken; replication of observations is therefore purely additive and is modelled as a grow-only set, whose merge is set union. Union is commutative, associative, and idempotent, so two replicas that have received the same set of observations hold the same state irrespective of the order in which they received them, of duplicate delivery, and of arbitrary delay. No conflict resolution arises for this class, and no observation can be discarded by merging, which matters because applying a last-writer-wins rule to observations would silently drop measurements recorded concurrently at two gateways. *Mutable state*, by contrast, does admit genuine conflicts: calibration coefficients, irrigation thresholds, node configuration, and pointers to deployed model versions may be updated at the edge and in the cloud independently, and the replicas must agree on one value. This second class is modelled as a map of last-writer-wins registers, one per key.

The second distinction concerns the merge rule for those registers, which must be a single total order if the type is to converge. Each versioned item carries the metadata tuple

```
[gateway_id, logical_clock, timestamp, value]
```

in which `gateway_id` uniquely identifies the originating gateway, `logical_clock` is a Lamport counter [54], `timestamp` is the wall-clock time at which the update was created, and `value` is the updated content. Writing $v = (c_v, t_v, g_v)$ for the ordering key of a version, the merge of two versions of the same register is

$$v_1 \sqcup v_2 = \begin{cases} v_1, & \text{if } (c_2, t_2, g_2) <_{\text{lex}} (c_1, t_1, g_1), \\ v_2, & \text{otherwise,} \end{cases} \tag{1}$$

where $<_{\text{lex}}$ is the lexicographic order on the triple and gateway identifiers are totally ordered. The logical clock is therefore decisive, the wall-clock timestamp breaks ties between updates with equal logical clocks, and the gateway identifier breaks the remaining ties deterministically. Because $<_{\text{lex}}$ is a total order on a finite tuple, the merge defined above is commutative, associative, and idempotent; it is a join on a join-semilattice, and the register consequently satisfies strong eventual consistency, meaning that replicas which have delivered the same set of updates hold equal state regardless of delivery order.

It is worth being explicit about why the ordering cannot be stated as a set of independent rules. A rule of the form "prefer the higher logical clock, except where the wall-clock timestamps differ by more than a threshold, in which case prefer the newer timestamp"

does not define a total order, because the exception makes the relation non-transitive. Merge then loses associativity, and two replicas that receive the same three updates in different orders can settle on different values, which is precisely the divergence the mechanism is meant to exclude. The threshold formulation is therefore not used; the wall-clock timestamp enters only as a subordinate tie-break within the single order above.

The logical clock is maintained under the usual Lamport discipline: a gateway increments its counter before each local update, and on merging a remote version it sets its counter to the maximum of the local and remote values [54]. This is what makes the ordering consistent with causality, so that an update which causally follows another cannot lose to it. A wall-clock timestamp could not supply that property on its own, because gateway clocks drift, and a real-time clock that is reset after a power failure can step backwards; bounded clock skew, obtained in practice from a GNSS receiver or from periodic time synchronisation, therefore affects only the tie-break between genuinely concurrent updates and never the ordering of causally related ones.

The cost of the register semantics should also be acknowledged. A last-writer-wins register converges, but it is lossy under concurrency: of two genuinely concurrent configuration changes, one is discarded rather than reconciled. That is acceptable for configuration, where the intent is that the most recent instruction prevails, and unacceptable for observations, which is the reason for the split described above. Where both the edge and the cloud may legitimately contribute to the same mutable collection, for example a set of farmer-supplied and cloud-generated labels, an observed-remove set is used instead of a register, so that concurrent additions all survive the merge.

Synchronisation itself uses the delta-state form of the above types [55]. Rather than shipping the full replica state, a gateway transmits join-irredundant delta-mutators accumulated since the last acknowledged exchange, and the receiver joins them into its own state. Deltas are exchanged under an anti-entropy protocol with acknowledged causal context; if an acknowledgement is lost, or if the buffer of unacknowledged deltas has been garbage-collected during a long outage, the protocol falls back to a full-state join, which is always safe because the join is idempotent. The metadata overhead is one ordering tuple per mutable key rather than per observation, which keeps the cost proportional to the configuration state rather than to the volume of telemetry. Delta synchronisation is expected to reduce network traffic relative to full-state transmission, but the magnitude of the reduction depends on the rate at which the state actually changes and would have to be measured in a deployed system rather than assumed.

Illustrative worked example (data-loss risk): To make the buffering risk concrete, consider a representative farm gateway with 12 connected sensors, each sending one measurement every 60 s. If the gateway loses power 6 h before it can synchronize with the central server, roughly $12 \times 60 \times 6 = 4320$ sensor messages would be buffered and potentially lost, a hypothetical scenario used only to size the problem, not a measurement from a deployed system. This risk is mitigated with a local buffering mechanism by an embedded database and a battery back-up system in the gateway. The battery back-up will supply about 20 min of operation time to allow the gateway to finish outstanding write operations, to synchronize the latest buffered data, and to perform a graceful shutdown. A small UPS or a solar-powered buffer is recommended for critical sensor installations, for example in remote agricultural stations or environmental monitoring points.

Turning to the processing performed at the gateway, it hosts a local time-series database, an inference engine for pre-trained machine learning models, and a rule-based automation controller. This configuration enables the following operations to proceed entirely without cloud connectivity:

- Real-time soil moisture monitoring and automatic irrigation triggering.

- Frost and heat stress alerts derived from local weather station data.
- Early disease-risk scoring based on leaf wetness and humidity thresholds.

Sensor readings, on the other hand, are buffered locally with conflict-free merge semantics. When an internet connection is re-established, the gateway synchronises queued data with the cloud platform using a prioritised, delta-based replication protocol intended to limit bandwidth consumption. The intent of this architecture is to preserve local operational continuity in low-connectivity environments while retaining the long-term analytical benefits of centralised cloud storage and model retraining. It should be read as an intent rather than as a guarantee: continuity holds for the operations listed above and for as long as the gateway itself has power and its buffer has capacity, and the failure modes discussed earlier in this subsection bound it. Whether the intent is met in practice is an empirical question, and Section 6.4 states how it would be tested.

5.4. Solution 4: Hybrid UAV Fleet Management Framework

The operational constraints of individual drones, including limited battery endurance, payload restrictions, and vulnerability to adverse weather, limit the spatial and temporal coverage achievable by single-unit deployments [4,29,30]. A Hybrid UAV Fleet Management Framework could be deployed, combining cooperative drone swarms with fixed ground-based sensing infrastructure to achieve continuous, scalable field coverage. Swarm coordination is handled by a distributed mission planner that dynamically partitions field scan tasks among available units based on battery state, sensor payload, and current wind conditions. Individual drones upload incremental observations to the edge gateway upon each return, enabling the system to construct a continuously updated field mosaic without requiring any single unit to complete a full survey flight.

A geofencing and regulatory compliance module could apply operator airspace restrictions directly in the flight planner, reducing the regulatory burden on farm operators [27,29]. Energy autonomy is extended through solar-assisted charging stations distributed at field boundaries, minimising turnaround time between sorties. The intended effect of this hybrid approach is to extend effective operational coverage beyond what a single unit can achieve, while remaining consistent with prevailing regulatory constraints on BVLOS operations. The magnitude of any extension depends on fleet size, field geometry, and charging-station placement, and would have to be established by the simulation and field protocol of Section 6.4 rather than asserted here.

5.5. Solution 5: Semantic Middleware for Data Interoperability and Standardization

To address the fragmentation caused by incompatible protocols, data formats, and vendor-specific APIs [19,22,46], we propose a Semantic Agricultural Middleware Layer positioned between heterogeneous field devices and upstream analytics platforms. The middleware performs real-time protocol translation across dominant communication standards including LoRaWAN, ZigBee, NB-IoT, Wi-Fi, and Bluetooth Low Energy (BLE) [17,18], normalising all incoming streams into a single canonical data model.

Data representation relies on established open standards: Message Queuing Telemetry Transport and Representational State Transfer APIs govern machine-to-machine communication, while JavaScript Object Notation for Linked Data (JSON-LD) and the Open Geospatial Consortium (OGC) SensorThings API provide ontology-aligned semantic context. Every data record is enriched with provenance metadata encompassing device identifier, geolocation coordinates, calibration timestamp, crop type, and measurement unit [8,11]. This provenance register enables downstream services to assess data quality and lineage without querying individual device vendors.

The proposed approach has three strengths. In terms of protocol abstraction, a single integration point replaces the point-to-point connectors that would otherwise be required between every device and platform pair. In terms of schema mapping and unit normalisation, automated transformation pipelines remove the need for manual data cleaning. In terms of vendor independence, a new sensor type is onboarded by registering a translation adapter, and the analytics layer is left unmodified [11,25].

The intended outcome is a plug-and-play integration fabric that lowers the engineering overhead of scaling heterogeneous smart-farming deployments. The size of that reduction is an empirical quantity, most directly expressed as the effort required to onboard a new device type, and it is not established by the design argument alone; Section 6.4 states how it would be measured.

5.6. Solution 6: Federated Agro-Learning Network for Labeled Data Scarcity

The scarcity of large, accurately labelled agricultural datasets constrains the generalisation capability of supervised machine learning models, particularly for rare disease events and crop-specific anomalies [27,34,48]. To address this, a Federated Agro-Learning Network could be deployed, in which model training is distributed across participating farms without requiring raw data to leave the farm premises. For example, each farm could train a local model on its own sensor observations and image captures, where only encrypted model weight updates (not the underlying data) are transmitted to a central aggregation server, and a federated averaging algorithm could synthesise a global model incorporating knowledge from all participants [7,34]. This approach reduces the exposure of raw farm data, since observations remain on the premises, while allowing a model to generalise across diverse agro-climatic conditions. It should not be described as preserving privacy without qualification. Model updates can themselves leak information about the data they were computed from, through membership-inference and reconstruction attacks, and federated averaging alone provides no formal guarantee against this. Where the threat model requires such a guarantee, the aggregation step has to be combined with secure aggregation, so that the server observes only the sum of updates, or with a differentially private training procedure that bounds the contribution of any single record, at a cost in model accuracy that has to be budgeted explicitly.

To address the specific challenge of label scarcity, the framework should incorporate three complementary strategies. First, self-supervised learning techniques would extract meaningful representations from unlabelled sensor streams, reducing dependence on annotated examples. Second, process-based agronomic simulation models, such as established crop growth and soil water models, would generate synthetic training data for rare events such as frost damage or pest outbreaks that are infrequently captured in real deployments [35,41]. Third, a lightweight human-in-the-loop mobile application would enable farmers and agronomists to contribute verified labels during routine field inspections, incrementally enriching the shared dataset over time.

Federated learning offers a privacy-preserving mechanism for collaborative model training across multiple farms; however, agricultural environments are inherently heterogeneous. Differences in crop types, soil properties, climate conditions, and management practices result in non-independent and identically distributed (non-IID) data, which can significantly degrade the performance of globally aggregated models. For example, a model trained jointly on wheat farms with loamy soils and rice farms with clay soils may generalize poorly to both environments.

To mitigate the non-IID challenge, the framework should incorporate techniques such as FedProx and stratified aggregation. For example, FedProx would introduce a proximal regularization term that would limit excessive divergence between local and global model

updates, reducing the performance gap and improving convergence stability, albeit at the cost of additional training rounds. Stratified aggregation would further improve performance by weighting model contributions according to similarities in crop type, soil characteristics, or climatic conditions rather than treating all farms equally.

Moreover, to address another limitation, the one of communication overhead over constrained network conditions against the tens of megabytes required by Large deep-learning models per synchronization cycle, the framework should implement compression and sparsification techniques. A further concern is the quality of the synthetic data generated by process-based simulation models. To mitigate this risk, the framework should deploy confidence-weighted training, in which synthetic samples receive lower importance than real observations. This preserves the benefits of synthetic data while limiting the influence of simulation error on the resulting model.

Finally, sustaining farmer participation in federated learning networks also requires appropriate incentive mechanisms. Since individual farms may contribute unequally to improvements in the global model, contribution-aware reward schemes are necessary to ensure fairness. The proposed framework employs Shapley value-based contribution assessment to quantify each farm's impact on model performance and allocate benefits accordingly. However, combining farms from substantially different environmental regions may introduce severe non-IID effects that reduce overall model performance. Therefore, regionalized federated learning deployments are generally recommended to maximize both predictive accuracy and participant engagement.

5.7. Solution 7: Explainable Decision Support for AI-Driven Recommendations

The opacity of deep learning models can reduce farmers' willingness to trust and act on automated recommendations, limiting the practical benefits of AI-driven agriculture [23,33,34]. To address this challenge, an Explainable Decision Support Dashboard could be applied. Every prediction would be augmented with human-readable explanations, confidence estimates, and the key factors that influenced the recommendation. The framework integrates Explainable Artificial Intelligence (XAI) techniques, combining SHAP and rule extraction methods to improve transparency and support informed decision-making [34,46].

SHAP decomposes model predictions into additive feature contributions, identifying which sensor measurements and environmental variables most strongly influence a given outcome. SHAP provides a consistent and theoretically grounded attribution of feature importance. Rule extraction further complements SHAP by translating model behavior into natural-language explanations that are easier for non-technical users to interpret. For example, in a hypothetical illustrative scenario, an irrigation recommendation of 35 mm could be accompanied by an explanation showing that a soil moisture level of 60% reduced the requirement, while a forecast temperature of 32 °C increased it because of higher expected evapotranspiration. Additional factors, such as rainfall probability and crop growth stage, are also presented to explain the final recommendation. Confidence intervals and data freshness indicators are displayed alongside the prediction to help farmers assess its reliability.

Although SHAP improves interpretability, practical deployment requires consideration of computational overhead and explanation stability. Generating SHAP values can be expensive for large-scale systems; therefore, caching previously computed explanations and using approximation techniques such as KernelSHAP can reduce execution costs. Since SHAP primarily provides local explanations, temporal comparison of explanations can also be used to clarify why recommendations change over time. Furthermore, uncer-

tainty indicators should accompany explanations when model confidence is low to prevent overreliance on potentially unstable predictions.

Finally, all recommendations should be presented through a multilingual interface in the farmer's preferred language. The purpose of providing transparent reasoning, confidence information, and intuitive explanations is to move farmers from being passive recipients of automated recommendations towards being informed decision-makers. Whether the dashboard has that effect is an empirical question about people rather than about the model, and it is not settled by the presence of the explanation mechanism. Explanations can also induce misplaced confidence in an incorrect recommendation, so the quantity of interest is appropriate reliance rather than trust as such. Section 6.4 states the user study through which this would be assessed.

5.8. Solution 8: Decentralised Identity Federation for Cross-Farm Access Control

Farming operations increasingly rely on shared physical assets—precision tractors, drone fleets, soil-sampling equipment—and multi-provider analytics ecosystems, creating complex trust relationships that conventional centralised identity management cannot adequately address [7,46,47]. For this purpose, a model for Identity Federation for Agriculture based on Decentralised Identifiers (DIDs) and Verifiable Credentials could be deployed. Each principal in the agricultural ecosystem, whether a machine, a permanent worker, a seasonal contractor, or a cloud service provider, would hold a self-sovereign DID anchored to a lightweight distributed ledger. Access to farm assets could be granted through time-bound, scope-limited authorisation tokens generated from verifiable credential presentations.

For example, a contractor borrowing a precision seeder for a single day receives a token valid for eight hours, restricted to that specific machine and the designated field boundary, and upon task completion the token is automatically revoked. Additionally, a delegation engine could allow farm managers to define hierarchical permission trees. This would enable cooperative managers or agronomists to sub-delegate limited access rights to third parties without involving the original data owner in every transaction. This architecture would support the emerging sharing economy for precision agriculture equipment while maintaining strict, auditable access boundaries [7,47]. One dependency should be noted, because it interacts with the offline-first design of Solution 3: verifying a credential presentation is possible offline, whereas learning that a credential has been revoked is not. A gateway that has been disconnected will therefore continue to accept a revoked token until it next reconciles, and the design bounds this by keeping token lifetimes short relative to the expected reconnection interval rather than by relying on immediate revocation.

5.9. Solution 9: Farmer-Centric Data Sovereignty Platform

Concerns over data misuse, opaque contractual terms, and uncompensated commercial exploitation remain primary barriers to data sharing in agricultural communities [7,46–48]. To address this concern, a Farmer-Centric Data Sovereignty Platform could be deployed, in which each farmer maintains exclusive control over their data through a personal, encrypted data wallet. Access policies are expressed as fine-grained permissions and enforced through smart contracts, providing automated, tamper-evident, and auditable governance without reliance on centralised intermediaries. Illustrative permission properties could include:

- Share anonymised yield statistics with the regional cooperative for 30 days;
- Share soil nutrient profiles with a specified agronomist, revocable on demand;
- Prohibit use of any farm data for third-party advertising or price-modelling purposes.

Access control is enforced through an attribute-based access control model, which applies these policies at the data layer so that even authorised recipients receive only the

attributes their permission scope covers. All data are encrypted at rest and in transit, and a cryptographically signed consent log records every access event, providing an audit trail intended to be sufficient for legal purposes; whether it is in fact sufficient depends on the jurisdiction and on the evidential standards applied there, which is one of the reasons the legal review recommended below is necessary. The intent of the architecture is to convert data sharing from an act of trust into a controlled and revocable transaction. Whether that intent translates into higher farmer participation is an adoption question that this work does not answer, and Section 6.4 states how it would be examined.

Considering the *Farmer Usability and Legal Validation*, a significant challenge in agricultural data sharing is that many farmers have limited digital literacy and may struggle to understand the legal and technical implications of smart contracts. To improve accessibility, the proposed framework replaces free-text agreements with simplified permission templates that clearly describe what data will be shared, who can access it, and for what purpose. The template-based approach described above has not yet been evaluated with farmers; no usability study, participant recruitment, or ethics review was conducted as part of this work. Based on general findings in the human–computer interaction and legal-design on simplified consent interfaces, we would expect structured permission templates to improve comprehension and completion time relative to free-text contracts, and to increase farmers' subjective comfort with data-sharing decisions. Confirming this expectation, including comprehension accuracy, completion time, and comparative confidence relative to conventional contracts, requires a dedicated usability study with an appropriate sample of farmers, informed consent, and, where applicable, institutional ethics approval; this is identified as a concrete direction for future empirical work, as it is discussed later under Section 7.

Although permission templates improve usability, privacy risks remain because even anonymized datasets may reveal sensitive information through metadata. For example, an attacker could combine publicly shared yield records with external information, such as land registries, to infer the identity of individual farmers. To mitigate this threat, the framework should incorporate differential privacy mechanisms. Laplace noise could, for example, be added to published yield values as:

$$(\text{Published} = \text{True_Yield} + \text{Lap}(0,\sigma)),$$

where (σ) can be set to limit the probability of successful re-identification.

The framework also considers regulatory compliance requirements. Under the General Data Protection Regulation, farmers are treated as data subjects, while cooperatives act as data controllers or processors. Smart contracts should therefore be designed so as to support consent withdrawal and Right-to-Be-Forgotten requirements by enabling farmers to revoke permissions through a single action, after which associated data are deleted within the required retention period. Audit logs provide verifiable evidence of deletion and consent management. However, despite these safeguards, several limitations remain, as questions regarding ownership and intellectual property rights for machine-learning models trained on shared agricultural data remain unresolved. Consequently, legal review is recommended before enabling large-scale cross-farm data sharing or collaborative machine-learning initiatives.

5.10. Solution 10: Smart Agriculture as a Service (SAaaS) for Inclusive Adoption

The substantial capital expenditure required for precision sensing hardware, automated machinery, and analytics platforms creates a prohibitive financial barrier for smallholder and resource-constrained farmers, risking a structural digital divide between large-scale industrial operations and subsistence-level producers [37,38,47,48]. For this

purpose a SAaaS model could be deployed to transform capital expenditure into affordable operational expenditure through a cooperative subscription framework.

Under this model, sensors could be leased on a monthly basis and remain the property of the service provider, who bears maintenance and replacement costs. Drone scouting services are offered on-demand at per-hectare pricing, enabling small farms to access aerial intelligence without owning a UAV. A shared data platform could aggregate anonymised observations from participating farms to improve model accuracy for all members, creating a network effect that increases service quality proportionally with subscriber growth [46,48].

The pricing structure integrates government subsidy mechanisms through direct deduction at point of subscription, reducing administrative friction for programme beneficiaries. A multilingual, low-bandwidth mobile application may provide the primary farmer interface, designed for operation on entry-level smartphones under intermittent connectivity conditions. Digital literacy training modules could be embedded within the application onboarding flow. By disaggregating the cost of adoption and aligning it with cash-flow cycles, the SAaaS model could lower the barrier to entry for smallholder farmers in both developed and developing agricultural economies [37,47]. This is a proposition about costs and incentives rather than a demonstrated result: it depends on subscriber numbers, on hardware failure rates, and on the availability of the subsidy mechanisms assumed above, and it would be tested by the techno-economic modelling described in Section 6.4.

Table 4 summarises the ten problem–solution pairs presented above and indicates the layer of the architecture of Section 6 at which each solution operates. The wording of the first column is identical to the wording used in Section 4, so that the two lists can be read against each other without ambiguity.

6. A Layered Trustworthy Framework for Smart Agriculture

The preceding solutions address individually distinct challenges. Implementing them in isolation, however, risks a fragmented system in which security, interoperability, connectivity resilience, explainability, and governance operate as disconnected silos. A *Layered Trustworthy Architecture* for smart agriculture is therefore proposed, integrating all ten solutions into a single end-to-end framework. Rather than stacking independent subsystems, the architecture treats security, provenance, and access control as cross-cutting concerns enforced at every layer.

This section synthesizes the individually proposed solutions into a single, layered architecture for trustworthy smart agriculture. The five architectural layers are discussed in turn: sensing and device, connectivity and edge, interoperability and data integration, intelligence and decision support, and governance, ownership, and trust. The end-to-end workflow that links them is then presented. The architecture has not been implemented or evaluated, and is offered as a conceptual synthesis that consolidates the solutions of Section 5 and identifies directions for future work. Table 4 maps each conceptual solution to the open problem it addresses and to the layer at which it operates.

6.1. Architecture Overview

The proposed system is organised into five tightly coupled layers, each encapsulating a distinct set of responsibilities while exposing well-defined interfaces to the adjacent layers: (L_1)–(L_5). Data originate at field devices, traverse local edge infrastructure for real-time processing and buffering, are normalised and enriched by the interoperability middleware, consumed by cloud-scale analytics services, and ultimately delivered to farmer-facing dashboards governed by consent-based access policies. Security, provenance registration, and identity verification are enforced continuously throughout this pipeline rather than at

a single checkpoint. Figure 2 illustrates the proposed unified architecture for trustworthy smart agriculture, covering the complete data life-cycle from field sensing to end-user decision-making.

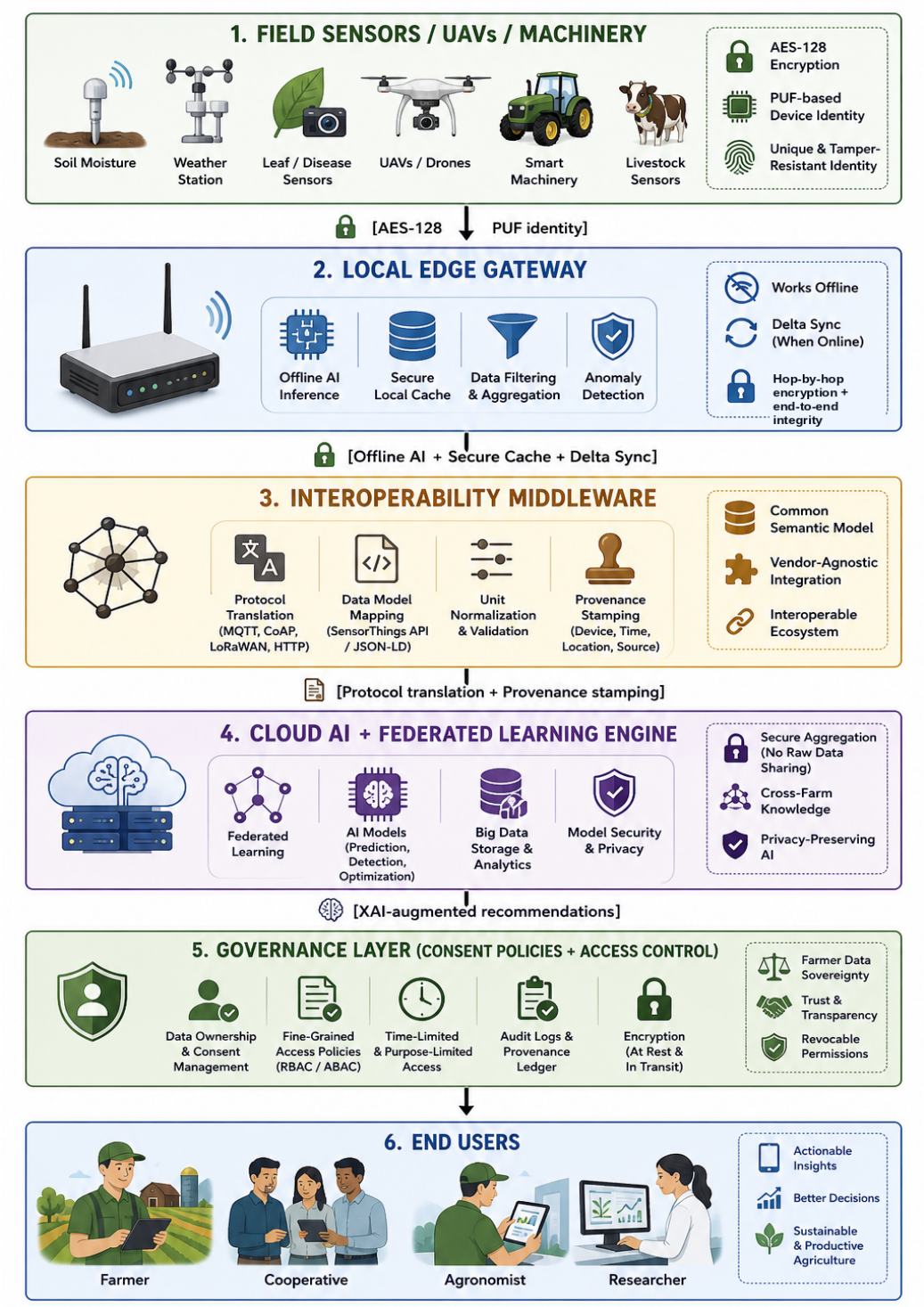


Figure 2. Unified architecture for trustworthy smart agriculture. The figure illustrates the end-to-end data pipeline from field sensing devices to end users, integrating edge computing, interoperability middleware, cloud-based AI with federated learning, explainable decision support, and farmer-centric data governance mechanisms.

The *Sensing and Device* layer (L_1), encompasses the heterogeneous field device ecosystem: soil moisture and nutrient sensors, microclimate weather stations, livestock biometric trackers, irrigation actuators, optical cameras, UAVs, and smart agricultural machinery.

Drawing on Solutions 1, 2, and 4, each device node is assigned a hardware-rooted digital identity via PUF-based device certificates and employs AES-128 symmetric encryption with rolling session keys for all transmissions. Modular sensor cartridge design and predictive maintenance algorithms extend operational lifetimes in harsh field conditions, while the UAV fleet coordination engine provides scalable aerial monitoring coverage.

In the *Connectivity and Edge* layer (L_2), an on-farm edge gateway serves as the local intelligence hub, collecting data through LoRaWAN, ZigBee, BLE, Wi-Fi, and NB-IoT links (Solution 3). Drawing on Solutions 1 and 2, the gateway manages certificate-based mutual authentication with sensor nodes, performs real-time anomaly filtering, executes predeployed machine learning inference models, and maintains a local buffered data store. Critical agricultural actions—irrigation control, frost alerts, disease-risk notifications—execute autonomously without cloud dependency. Data are queued for asynchronous synchronisation, transmitted to the cloud using delta-based replication when connectivity permits, and reconciled using conflict-free merge semantics.

The *Interoperability and Data Integration* layer (L_3) implements Solution 5, operating as a semantic middleware service that receives data from one or more farm edge gateways and transforms heterogeneous payloads into a unified canonical model conforming to the OGC SensorThings API and JSON-LD agricultural ontologies. Operations include protocol translation, schema mapping, unit normalisation, metadata enrichment, and provenance record creation. Each processed record is stamped with a unique data lineage identifier linking it to its originating device, gateway, and farm, establishing an append-only audit trail consumed by the governance layer. The identifier also records which attestation the record carries, in the sense of Solution 2: whether it is a node-attested measurement or a gateway-attested derivative. The trail is append-only rather than immutable in the strict sense, since immutability would additionally require that the storage medium itself be tamper-proof, which is an operational property of the deployment and not of the middleware.

Through *Intelligence and Decision Support* layer (L_4), cloud-hosted and edge-resident AI services provide yield forecasting, disease detection, irrigation scheduling, resource optimisation, and machinery prognostics (Solution 6). The federated learning engine periodically aggregates encrypted model updates from participating farms, continuously improving global model accuracy without exposing raw farm data. Every recommendation delivered to the farmer interface is generated through the Explainable AI pipeline (Solution 7): SHAP-based feature attribution, natural-language rule extraction, and calibrated confidence intervals accompany each prediction. Dashboards adapt presentation complexity to farmer experience level and render outputs in the user's preferred language, supporting informed human oversight of all automated recommendations.

Finally, the *Governance, Ownership, and Trust* layer (L_5), spans the entire data lifecycle, implementing farmer-centric data sovereignty (Solution 9) and federated identity management (Solution 8). Each farmer maintains a personal encrypted data wallet governed by smart-contract access policies that specify which data attributes may be shared, with whom, for what duration, and for what declared purpose. Attribute-based access control enforces these policies at the data plane; time-limited, scope-restricted tokens govern cross-farm equipment sharing and contractor access. A cryptographically signed provenance ledger records every data access event, with the intention of providing an auditable chain of custody from sensor to end consumer. Its completeness is conditional: it covers accesses mediated by the platform, and it cannot record what a recipient does with data after a permitted disclosure, which remains a contractual rather than a technical control.

6.2. End-to-End Architecture Workflow

The workflow of the proposed *Layered Trustworthy Architecture* for smart agriculture is presented below. It follows the data through each of the five incorporated layers, namely the *Sensing and Device* layer, the *Connectivity and Edge* layer, the *Interoperability and Data Integration* layer, the *Intelligence and Decision Support* layer, and finally the *Governance, Ownership, and Trust* layer. The five layers are followed by a sixth step, which is not a layer of the architecture but the delivery of its outputs to the people who act on them; it is listed here because Figure 2 shows it as the final block of the pipeline. The names used in the steps below are the names used for the layers in Section 6 and in Table 4.

1. The workflow begins at the *Sensing and Device* layer (L_1), which includes heterogeneous data sources such as field sensors, UAVs, and smart agricultural machinery. These devices collect environmental, crop, and operational data and secure it using lightweight cryptographic mechanisms, including symmetric encryption (e.g., AES-128) and hardware-based identities such as PUFs.
2. Data are then transmitted to the *Connectivity and Edge* layer (L_2), realised as the on-farm edge gateway, where an edge-first processing model is applied. This layer performs offline AI inference, temporary secure storage (cache), and anomaly detection. Importantly, it may enable continuous operation even in the absence of internet connectivity by supporting local decision-making and delayed synchronization with cloud services when connectivity is restored.
3. Next, the data enter the *Interoperability and Data Integration* layer (L_3), where heterogeneity is addressed through protocol translation, data model harmonization, and unit normalization. Additionally, provenance metadata are attached to each data stream, ensuring traceability by recording information such as data source, timestamp, and location.
4. The processed data are then forwarded to the *Intelligence and Decision Support* layer (L_4), hosted in the cloud and, for inference, also at the edge. This layer provides scalable storage and advanced analytics, including machine learning and cross-farm knowledge sharing through federated learning techniques. To enhance trust and usability, the system integrates XAI, generating transparent and interpretable recommendations for agricultural decision-making.
5. The *Governance, Ownership, and Trust* layer (L_5) enforces data ownership, privacy, and access control. Farmers retain sovereignty over their data through fine-grained consent policies, time-limited access permissions, and auditable data usage logs. Encryption is applied both at rest and in transit. Consistently with the trust boundary set out in Solution 2 (see, Section 5.2), this provides confidentiality hop by hop, with the edge gateway inside the boundary, together with integrity and origin authentication that remain verifiable along the whole path from the originating node.
6. Finally, and outside the five layers proper, insights are delivered to end users, including farmers, cooperatives, agronomists, and researchers. These stakeholders interact with the system through intuitive dashboards that provide actionable recommendations, enabling informed, transparent, and sustainable agricultural practices. The architecture integrates security, interoperability, resilience, explainability, and governance into a unified framework, addressing key challenges in smart agriculture systems.

6.3. Positioning with Respect to Existing Smart-Agriculture Architectures

None of the components assembled above is new. PUF-based device identity, authenticated symmetric telemetry, replicated data types, semantic middleware over the OGC SensorThings API, federated averaging, SHAP-based attribution, decentralised identifiers, and subscription service models all pre-exist this work, and the review of Section 2 cites

the literature in which they were developed. Any claim to contribution must therefore rest on the composition rather than on the parts, and it is worth stating precisely what that composition is asserted to add, so that the assertion can be disputed or tested.

Table 5 compares the proposed architecture with the principal architectural families identified in Sections 2.4 and 2.9. The comparison records where each family places its emphasis, judged by what the cited works specify and evaluate. It is not a claim that these architectures could not be extended in the remaining directions; most could be, and several of the components used here originate in them. The point of the table is that the reviewed architectures address subsets of the trust properties, and that the subsets differ.

Table 5. Coverage of trust-related properties by the principal smart-agriculture architectural families and by the architecture proposed here. Entries record the emphasis of the cited works, not an upper bound on what those designs could accommodate. "Partial" indicates that the property is addressed but is not a design constraint on the architecture as a whole.

Architectural Family	Hardware-Rooted Device Identity	Offline Operation with Defined Merge Semantics	Semantic Interoperability with Provenance	Explainable Decision Support	Farmer-Held Consent and Access Control	Cost and Inclusion
Cloud-centric IoT [20,25]	No	No	Partial	No	No	No
Layered WSN/IoT reference models [16,17]	Partial	No	No	No	No	No
Edge-first and AIoT designs [10,22,23,26]	Partial	Partial	No	Partial	No	No
Federated and data-space oriented [7,46]	Partial	No	Partial	No	Partial	No
Data-engineering and IoRT architectures [11]	No	No	Yes	No	Partial	No
Proposed (Section 6)	Yes	Yes	Yes	Yes	Yes	Yes

Against that background, the integration is claimed to add only the following:

1. **A single provenance spine.** Device identity, lineage metadata, and consent enforcement are usually specified in separate architectures, with separate identifiers. Here they are chained: the hardware-rooted identity of Solution 2 (see, Section 5.2) is what the lineage identifier of Solution 5 (see, Section 5.5) attests to, and that same identifier is the subject of the consent records of Solution 9 (see, Section 5.9). A consequence is that a governance decision can be traced to the physical device that produced the datum, and an attested measurement can be distinguished from a gateway-derived aggregate, which neither a provenance-only nor an identity-only design supports on its own.
2. **Trust properties as cross-layer invariants.** Security, provenance, and access control are stated as properties that every layer must preserve and hand on, with the interfaces between layers defined in those terms, rather than as functions assigned to particular layers. The trust-boundary analysis of Solution 2 (see, Section 5.2) is an instance: it is meaningful only because the layer above it agrees to carry the node attestation rather than to replace it.
3. **Intermittent connectivity as a constraint on every layer, including governance.** Offline capability is commonly a property of the edge layer alone. Here it also constrains the governance layer, which is why the revocation-latency dependency is stated explicitly in Solution 8 (see, Section 5.8) rather than assumed away, and it constrains the data layer, which is why the replicated types of Solution 3 (see, Section 5.3) are specified as convergent rather than as best-effort buffering.
4. **Inclusion of the socio-economic layer in the architecture.** Cost, digital literacy, and the terms of participation are treated as an architectural layer with interfaces to the

technical layers, rather than as adoption commentary appended to a technical design. Problems 9 and 10 are answered inside the architecture, not outside it.

These four claims are argued rather than demonstrated. The argument establishes at most that the composition is coherent and that it covers properties the reviewed architectures cover separately; it does not establish that the composition performs better than the alternatives, or that the coupling between layers is worth the integration cost it imposes. Establishing either would require the evaluation described next, and in particular a comparison against a baseline in which the same components are deployed without the cross-layer coupling. Until that comparison exists, the incremental value stated here should be read as the hypothesis the architecture is designed to test.

6.4. Evaluation Plan for Future Validation

No component of the architecture has been implemented or measured, thus the claims made for it are testable propositions rather than results. Setting out how each would be tested is part of the contribution: it fixes the metrics in advance, names the baselines against which any improvement would have to be shown, and makes clear which claims are engineering questions and which are questions about people. Table 6 states the plan component by component, following the numbering of Section 5.

Table 6. Evaluation plan for the ten conceptual solutions. For each solution the table states the method by which it would be validated, the testbed or data required, the primary metrics, and the baseline against which any claimed improvement would have to be demonstrated. None of these evaluations has been carried out in this work.

#	Solution	Method	Testbed or Data	Primary Metrics	Baseline
1	Self-healing modular nodes	Accelerated environmental ageing and multi-season field deployment	Climatic chamber (thermal cycling, humidity, UV); two crop sites over two seasons	Node survival rate; calibration drift over time; mean interval between maintenance visits; cost per node-year	Sealed non-modular nodes of equivalent sensing capability
2	Hierarchical edge security	Symbolic protocol verification under the Dolev–Yao model, plus microcontroller benchmarking	Protocol model in a verification tool; representative low-power microcontroller with PUF	Secrecy and authentication goals discharged; energy per authenticated message; added latency; PUF bit-error rate across temperature	DTLS with certificate-based PKI; pre-shared symmetric keys
3	Offline-first edge/fog architecture	Discrete-event simulation with injected outages, plus a two-gateway testbed	Replayed telemetry traces; scripted partition and clock-skew scenarios	Convergence after anti-entropy (state equality); lost or overwritten updates; bytes per synchronisation cycle; control-loop availability during outage	Full-state synchronisation; last-writer-wins on wall-clock time alone
4	Hybrid UAV fleet management	Multi-UAV mission simulation followed by a limited field trial	Simulated field geometries; small fleet with charging stations	Hectares surveyed per battery-hour; coverage refresh interval; mission completion rate under wind	Single-UAV sequential survey
5	Semantic middleware	Integration study across heterogeneous device types	At least five device families across three protocols	Effort to onboard a new device type (person-hours); mapping error rate; provenance field completeness	Point-to-point vendor connectors
6	Federated agro-learning network	Cross-silo federated experiments under controlled non-IID partitions	Public agricultural image and time-series datasets, partitioned by crop and soil	Global and per-site accuracy; rounds to convergence; communication volume per round; stability of Shapley contribution scores	Centralised training; plain federated averaging

Table 6. *Cont.*

#	Solution	Method	Testbed or Data	Primary Metrics	Baseline
7	Explainable decision support	Controlled user study with farmers and agronomists	Recommendation scenarios with known ground truth, including deliberately incorrect recommendations	Appropriate reliance (acceptance of correct and rejection of incorrect advice); comprehension; time to decision; explanation latency	Recommendation presented without explanation
8	Decentralised identity federation	Security analysis and prototype implementation	DID and verifiable-credential libraries; simulated intermittent connectivity	Token issuance and revocation latency; offline verification success rate; revocation propagation delay	Centralised role-based access control
9	Farmer-centric data sovereignty	Usability study with permission templates, plus independent legal review	Recruited farmer sample with informed consent and ethics approval; jurisdiction-specific legal assessment	Comprehension accuracy; completion time; successful revocation rate; re-identification risk at a stated privacy budget	Free-text data-sharing agreements
10	Smart Agriculture as a Service	Techno-economic modelling followed by a cooperative pilot	Cost data from a cooperative; subsidy parameters	Total cost of ownership per hectare; break-even subscriber count; adoption and retention rates	Outright purchase of equivalent equipment

Two evaluations concern the architecture as a whole rather than any single component, and neither is reducible to the rows of Table 6. The first is end-to-end provenance traceability, measured as the proportion of delivered recommendations for which every contributing measurement can be traced to an attesting device, together with the proportion that can be distinguished as node-attested rather than gateway-derived. The second is the integration cost implied by the cross-layer coupling described in Section 6.3, measured against a baseline in which the same components are deployed without that coupling. This second comparison is the one that would decide whether the claimed incremental value is real, and it is the evaluation the authors regard as the most important next step.

The evaluations differ in kind, and the difference is worth marking. Those for Solutions 2, 3, and 8 (see, Sections 5.2, 5.3, and 5.8, respectively) are largely formal or instrumented, and could be completed without field access. Those for Solutions 1 and 4 (see, Sectinos 5.1 and 5.4) require seasons of field time. Those for Solutions 7, 9, and 10 (see, Sections 5.7, 5.9, and 5.10, respectively) are studies of people and institutions, requiring recruitment, informed consent, and ethics approval, and they cannot be substituted by simulation. Prototyping in the first group is therefore the most direct next step, and it is the one against which the technical claims of Section 5 would first be tested.

7. Conclusions and Future Directions

This paper contributes a trust-oriented perspective on smart-agriculture sensor data and sensor data transmission network ecosystems in the context of the 4IR. Through a synthesis of the literature, ten recurring open problems were identified and linked to a unified layered architecture that integrates interoperability, security, explainability, connectivity resilience, provenance, and governance. The framework emphasizes trustworthy sensor data as a foundational requirement for reliable decision-making throughout the agricultural data lifecycle.

7.1. Potentials and Limitations

This paper combines two distinct contributions. Sections 2 to 4 synthesise the published literature into an overview of sensor data and sensor data transmission networks in smart agriculture, and into a set of ten open problems. That part rests on the cited sources and is subject to the usual limitations of a narrative, non-systematic review. In particular, and as stated in Section 2.1, the selection of studies reflects the judgement of the

authors, screening was not independently duplicated, and no claim of exhaustive coverage is made. Sections 5 and 6 then propose a conceptual layered architecture and component-level design patterns intended to address those open problems. These proposals have not been prototyped, deployed, or empirically evaluated: no field trial, simulation, or user study was conducted. The proposed architecture should be considered as a set of testable design hypotheses, rather than as validated engineering results. Three consequences of that positioning are stated in the body of the paper rather than left to the reader. The derivation of the ten open problems is given as an explicit four-step procedure with retention criteria in Section 4.1, and Table 3 traces each problem back to the reviewed works that motivate it, so that the foundation the solutions rest on can be inspected and disputed. What the integration is claimed to add over existing smart-agriculture architectures is stated as four specific claims in Section 6.3, together with the acknowledgement that those claims are argued rather than demonstrated. The means by which each component would be validated, including the metrics and the baselines, are set out in Section 6.4. Prototyping individual components and evaluating them against real or realistic agricultural data is the most direct next step and is left for future work. Within that step, the formal and instrumented evaluations, namely protocol verification for the security design, partition-and-convergence testing for the replicated data types, and prototype measurement of credential issuance and revocation, are the ones that can be completed without field access and are therefore the ones the authors intend to undertake first.

Two limitations of the derivation itself should be carried forward alongside the limitations of the review. The extraction and coding of difficulty statements were performed by the author team without independent duplication, so the boundaries between the ten categories reflect the authors' reading of the corpus. The retention threshold of three independent works is a convention rather than a validated criterion, and a different threshold would most plausibly admit further problems rather than remove any of these ten.

Despite rapid progress, multiple research gaps remain before agricultural CPS can scale safely and equitably. Model validity for living systems remains a core challenge, because the biological variability and open-system dynamics of agriculture make verification and validation harder than in engineered environments [35,41]. A second gap concerns edge–cloud partitioning, which still lacks standardised design rules. Reported deployments demonstrate measurable latency reductions from moving computation towards the edge, but generalisable methodologies for deciding what must run on the device, at the edge, or in the cloud, under constraints of energy, bandwidth, and safety, are still emerging [10,22,23].

7.2. Future Directions

Trustworthy AI and evidence of benefit are the principal adoption bottlenecks. The literature proposes AI and machine learning widely, but farmers and policymakers need credible, context-specific evidence of agronomic and economic benefit, and this evidence gap is itself identified as a barrier to adoption [34,46,47]. Privacy-preserving collaboration is promising but not solved. Federated learning can reduce centralised data exposure, yet real deployments must handle device and data heterogeneity, governance incentives, and adversarial behaviour such as poisoning and inference attacks, particularly in multi-stakeholder ecosystems [6,7]. Finally, interoperability and governance remain as important as the technology itself. Architectural alignment, common information models, and engagement with standards bodies are decisive for scaling beyond pilots. Future research must therefore integrate systems engineering, standardisation, and socio-economic incentives rather than focusing only on algorithms or sensors [17,25,48]. The convergence of terrestrial and non-terrestrial connectivity is likely to reshape rural coverage assumptions over

time, but agriculture-relevant performance, including latency variability, intermittency, and power cost, must be characterised with farm workload models rather than with generic telecom key performance indicators [18,21].

Author Contributions: Conceptualization, A.I.M. and I.P.; methodology, A.I.M. and I.P.; Writing—original draft, A.I.M., I.P., C.S., R.K. and I.K.; Writing—review & editing, A.I.M., I.P., C.S., R.K. and I.K.; Supervision, I.P. and I.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are contained within the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors used Claude Sonnet 5.5 (Anthropic), accessed October 2026, for AI-assisted copy editing of the manuscript and for the generation of figures. For copy editing, the tool was employed to improve readability, grammar, spelling, punctuation, and phrasing of text that was originally written by the authors. For figures, Claude was used to produce images by arranging pre-designed shapes (e.g., geometric primitives such as rectangles, circles, arrows, and connectors) into layouts specified by the authors; it did not create original artwork or depict novel data. No AI-generated scientific content, data analysis, interpretations, conclusions, or research findings were produced by the tool. All revisions and figures were reviewed, verified, and approved by the authors, who retain full responsibility for the accuracy, originality, and integrity of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Uztürk, D.; Büyüközkan, G. Industry 4.0 technologies in smart agriculture: A review and a Technology Assessment Model proposition. *Technol. Forecast. Soc. Change* **2024**, *208*, 123640. [CrossRef]
2. Padhiary, M.; Kumar, A.; Sethi, L.N. Emerging technologies for smart and sustainable precision agriculture. *Discov. Robot.* **2025**, *1*, 6. [CrossRef]
3. Gupta, S.; Chowdhury, S.; Govindaraj, R.; Amesho, K.T.; Shangdiar, S.; Kadhila, T.; Iikela, S. Smart agriculture using IoT for automated irrigation, water and energy efficiency. *Smart Agric. Technol.* **2025**, *12*, 101081. [CrossRef]
4. Maddikunta, P.K.R.; Hakak, S.; Alazab, M.; Bhattacharya, S.; Gadekallu, T.R.; Khan, W.Z.; Pham, Q.-V. Unmanned aerial vehicles in smart agriculture: Applications, requirements, and challenges. *IEEE Sens. J.* **2021**, *21*, 17608–17619. [CrossRef]
5. Pasca, E.M.; Delinschi, D.; Erdei, R.; Baraian, I.; Matei, O.D. A vulnerable-by-design IoT sensor framework for cybersecurity in smart agriculture. *Agriculture* **2025**, *15*, 1253. [CrossRef]
6. Zanella, A.R.A.; da Silva, E.; Albini, L.C.P. Security challenges to smart agriculture: Current state, key issues, and future directions. *Array* **2020**, *8*, 100048. [CrossRef]
7. Gupta, M.; Abdelsalam, M.; Khorsandroo, S.; Mittal, S. Security and privacy in smart farming: Challenges and opportunities. *IEEE Access* **2020**, *8*, 34564–34584. [CrossRef]
8. Bissadu, K.D.; Hossain, G.; Velagala, L.P. Identifying sensor data integrity threats of smart agriculture: A collaborative filtering approach. *Appl. Eng. Agric.* **2024**, *40*, 565–575. [CrossRef]
9. Zou, X.; Liu, W.; Huo, Z.; Wang, S.; Chen, Z.; Xin, C.; Bai, Y.; Liang, Z.; Gong, Y.; Qian, Y.; et al. Current status and prospects of research on sensor fault diagnosis of agricultural Internet of Things. *Sensors* **2023**, *23*, 2528. [CrossRef] [PubMed]
10. Shi, W.; Cao, J.; Zhang, Q.; Li, Y.; Xu, L. Edge computing: Vision and challenges. *IEEE Internet Things J.* **2016**, *3*, 637–646. [CrossRef]
11. Bimonte, S.; Bellocchi, G.; Pinet, F.; Chalhoub, G.; Sakr, M.; Skrzypczyński, P. Data engineering for sustainable agriculture: Developments, challenges, and case studies of a novel IoRT architecture. *J. Big Data* **2025**, *12*, 195. [CrossRef]
12. Ahmed, Z.; Gui, D.; Murtaza, G.; Yunfei, L.; Ali, S. An overview of smart irrigation management for improving water productivity under climate change in drylands. *Agronomy* **2023**, *13*, 2113. [CrossRef]
13. Aarif, K.O.; Alam, A.; Hotak, Y. Smart sensor technologies shaping the future of precision agriculture: Recent advances and future outlooks. *J. Sens.* **2025**, *2025*, 2460098. [CrossRef]
14. Yadav, A.; Yadav, K.; Ahmad, R.; Abd-Elsalam, K.A. Emerging frontiers in nanotechnology for precision agriculture: Advancements, hurdles and prospects. *Agrochemicals* **2023**, *2*, 220–256. [CrossRef]
15. Soussi, A.; Zero, E.; Sacile, R.; Trinchero, D.; Fossa, M. Smart sensors and smart data for precision agriculture: A review. *Sensors* **2024**, *24*, 2647. [CrossRef] [PubMed]

16. Ayaz, M.; Ammad-Uddin, M.; Sharif, Z.; Mansour, A.; Aggoune, E.-H.M. Internet-of-Things (IoT)-based smart agriculture: Toward making the fields talk. *IEEE Access* **2019**, *7*, 129551–129583. [CrossRef]
17. Mowla, M.N.; Mowla, N.; Shah, A.F.M.S.; Rabie, K.M.; Shongwe, T. Internet of Things and wireless sensor networks for smart agriculture applications: A survey. *IEEE Access* **2023**, *11*, 145813–145852. [CrossRef]
18. Pagano, A.; Croce, D.; Tinnirello, I.; Vitale, G. A survey on LoRa for smart agriculture: Current trends and future perspectives. *IEEE Internet Things J.* **2023**, *10*, 3664–3679. [CrossRef]
19. Aldhaheri, L.; Alshehhi, N.; Manzil, I.I.J.; Khalil, R.A.; Javaid, S.; Saeed, N.; Alouini, M.-S. LoRa communication for agriculture 4.0: Opportunities, challenges, and future directions. *IEEE Internet Things J.* **2025**, *12*, 1380–1407. [CrossRef]
20. Sanjeevi, P.; Prasanna, S.; Siva Kumar, B.; Gunasekaran, G.; Alagiri, I.; Vijay Anand, R. Precision agriculture and farming using Internet of Things based on wireless sensor network. *Trans. Emerg. Telecommun. Technol.* **2020**, *31*, e3978. [CrossRef]
21. Polymeni, S.; Plastras, S.; Skoutas, D.N.; Kormentzas, G.; Skianis, C. The impact of 6G-IoT technologies on the development of agriculture 5.0: A review. *Electronics* **2023**, *12*, 2651. [CrossRef]
22. Qazi, S.; Khawaja, B.A.; Farooq, Q.U. IoT-equipped and AI-enabled next generation smart agriculture: A critical review, current challenges and future trends. *IEEE Access* **2022**, *10*, 21219–21235. [CrossRef]
23. Mansoor, S.; Iqbal, S.; Popescu, S.M.; Kim, S.L.; Chung, Y.S.; Baek, J.-H. Integration of smart sensors and IoT in precision agriculture: Trends, challenges and future prospectives. *Front. Plant Sci.* **2025**, *16*, 1587869. [CrossRef] [PubMed]
24. Kassim, M.R.M. IoT applications in smart agriculture: Issues and challenges. In Proceedings of the 2020 IEEE Conference on Open Systems (ICOS), Kota Kinabalu, Malaysia, 17–19 November 2020; pp. 19–24. [CrossRef]
25. Araújo, S.O.; Peres, R.S.; Barata, J.; Lidon, F.; Ramalho, J.C. Characterising the agriculture 4.0 landscape—Emerging trends, challenges and opportunities. *Agronomy* **2021**, *11*, 667. [CrossRef]
26. Adli, H.K.; Remli, M.A.; Wan Salihin Wong, K.N.S.; Ismail, N.A.; González-Briones, A.; Corchado, J.M.; Mohamad, M.S. Recent advancements and challenges of AIoT application in smart agriculture: A review. *Sensors* **2023**, *23*, 3752. [CrossRef] [PubMed]
27. Phang, S.K.; Chiang, T.H.A.; Happonen, A.; Chang, M.M.L. From satellite to UAV-based remote sensing: A review on precision agriculture. *IEEE Access* **2023**, *11*, 127057–127076. [CrossRef]
28. Khanal, S.; Kc, K.; Fulton, J.P.; Shearer, S.; Ozkan, E. Remote sensing in agriculture—Accomplishments, limitations, and opportunities. *Remote Sens.* **2020**, *12*, 3783. [CrossRef]
29. Guebsi, R.; Mami, S.; Chokmani, K. Drones in precision agriculture: A comprehensive review of applications, technologies, and challenges. *Drones* **2024**, *8*, 686. [CrossRef]
30. Delavarpour, N.; Koparan, C.; Nowatzki, J.; Bajwa, S.; Sun, X. A technical study on UAV characteristics for precision agriculture applications and associated practical challenges. *Remote Sens.* **2021**, *13*, 1204. [CrossRef]
31. Aslan, M.F.; Durdu, A.; Sabanci, K.; Ropelewska, E.; Gültekin, S.S. A comprehensive survey of the recent studies with UAV for precision agriculture in open fields and greenhouses. *Appl. Sci.* **2022**, *12*, 1047. [CrossRef]
32. Bhat, S.A.; Huang, N.-F. Big data and AI revolution in precision agriculture: Survey and challenges. *IEEE Access* **2021**, *9*, 110209–110222. [CrossRef]
33. Garg, D.; Alam, M. Smart agriculture: A literature review. *J. Manag. Anal.* **2023**, *10*, 675–703. [CrossRef]
34. Linaza, M.T.; Posada, J.; Bund, J.; Eisert, P.; Quartulli, M.; Döllner, J.; Pagani, A.; Olaizola, I.G.; Barriguinha, A.; Moysiadis, T.; et al. Data-driven artificial intelligence applications for sustainable precision agriculture. *Agronomy* **2021**, *11*, 1227. [CrossRef]
35. Zhai, Z.; Martínez, J.F.; Beltran, V.; Martínez, N.L. Decision support systems for agriculture 4.0: Survey and challenges. *Comput. Electron. Agric.* **2020**, *170*, 105256. [CrossRef]
36. Vinod Chandra, S.S.; Anand Hareendran, S.; Albaaji, G.F. Precision farming for sustainability: An agricultural intelligence model. *Comput. Electron. Agric.* **2024**, *226*, 109386. [CrossRef]
37. Karunathilake, E.M.B.M.; Le, A.T.; Heo, S.; Chung, Y.S.; Mansoor, S. The path to smart farming: Innovations and opportunities in precision agriculture. *Agriculture* **2023**, *13*, 1593. [CrossRef]
38. Sharma, K.; Shivandu, S.K. Integrating artificial intelligence and Internet of Things (IoT) for enhanced crop monitoring and management in precision agriculture. *Sens. Int.* **2024**, *5*, 100292. [CrossRef]
39. Alahmad, T.; Neményi, M.; Nyéki, A. Applying IoT sensors and big data to improve precision crop production: A review. *Agronomy* **2023**, *13*, 2603. [CrossRef]
40. Monteiro, A.; Santos, S.; Gonçalves, P. Precision agriculture for crop and livestock farming—Brief review. *Animals* **2021**, *11*, 2345. [CrossRef] [PubMed]
41. Saiz-Rubio, V.; Rovira-Más, F. From smart farming towards agriculture 5.0: A review on crop data management. *Agronomy* **2020**, *10*, 207. [CrossRef]
42. Hayat, M.K.; Khan, T.Y.; Sial, T.N.; Khan, H.A.; Ashraf, Z.U.; Sheraz, M.; Oumar, B.A.; Usman, M.; Nadeem, M.; Belizaire, P.J. The biotech and precision agriculture revolution: Feeding the world sustainably. *Policy Res. J.* **2025**, *3*, 293–316. [CrossRef]

43. Sott, M.K.; Furstenau, L.B.; Kipper, L.M.; Giraldo, F.D.; Lopez-Robles, J.R.; Cobo, M.J.; Zahid, A.; Abbasi, Q.H.; Imran, M.A. Precision techniques and agriculture 4.0 technologies to promote sustainability in the coffee sector: State of the art, challenges, and future trends. *IEEE Access* **2020**, *8*, 149854–149867. [CrossRef]
44. Chapman, J.; Power, A.; Netzel, M.E.; Sultanbawa, Y.; Smyth, H.E.; Truong, V.K.; Cozzolino, D. Challenges and opportunities of the fourth revolution: A brief insight into the future of food. *Crit. Rev. Food Sci. Nutr.* **2022**, *62*, 2845–2853. [CrossRef] [PubMed]
45. Hassoun, A.; Aït-Kaddour, A.; Abu-Mahfouz, A.M.; Rathod, N.B.; Bader, F.; Barba, F.J.; Biancolillo, A.; Cropotova, J.; Galanakis, C.M.; Jambrak, A.R.; et al. The fourth industrial revolution in the food industry—Part I: Industry 4.0 technologies. *Crit. Rev. Food Sci. Nutr.* **2023**, *63*, 6547–6563. [CrossRef] [PubMed]
46. Jakku, E.; Taylor, B.; Fleming, A.; Mason, C.; Fielke, S.; Sounness, C.; Thorburn, P. "If they don't tell us what they do with it, why would we trust them?" Trust, transparency and benefit-sharing in smart farming. *NJAS—Wagening. J. Life Sci.* **2019**, *90–91*, 100285. [CrossRef]
47. Ahmed, B.; Shabbir, H.; Naqvi, S.R.; Peng, L. Smart agriculture: Current state, opportunities, and challenges. *IEEE Access* **2024**, *12*, 144456–144478. [CrossRef]
48. Vijayakumar, S.; Murugaiyan, V.; Ilakkiya, S.; Kumar, V.; Meenakshi Sundaram, R.; Mahender Kumar, R. Opportunities, challenges, and interventions for agriculture 4.0 adoption. *Discov. Food* **2025**, *5*, 265. [CrossRef]
49. Sharma, S.; Sharma, C.; Asenso, E.; Sharma, K. Research constituents and trends in smart farming: An analytical retrospection from the lens of text mining. *J. Sens.* **2023**, *2023*, 6916213. [CrossRef]
50. Dolev, D.; Yao, A. On the security of public key protocols. *IEEE Trans. Inf. Theory* **1983**, *29*, 198–208. [CrossRef]
51. Dworkin, M. *Recommendation for Block Cipher Modes of Operation: Methods and Techniques*; NIST Special Publication 800-38A; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2001. [CrossRef]
52. Bellare, M.; Canetti, R.; Krawczyk, H. Keying hash functions for message authentication. In *Advances in Cryptology—CRYPTO '96*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 1996; Volume 1109, pp. 1–15. [CrossRef] [PubMed]
53. Shapiro, M.; Preguiça, N.; Baquero, C.; Zawirski, M. Conflict-free replicated data types. In *Stabilization, Safety, and Security of Distributed Systems (SSS 2011)*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2011; Volume 6976, pp. 386–400. [CrossRef]
54. Lamport, L. Time, clocks, and the ordering of events in a distributed system. *Commun. ACM* **1978**, *21*, 558–565. [CrossRef]
55. Almeida, P.S.; Shoker, A.; Baquero, C. Delta state replicated data types. *J. Parallel Distrib. Comput.* **2018**, *111*, 162–173. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.