Federated Machine Learning Frameworks for Secure Agricultural Decision Support Systems: A Conceptual Review

Nidhi Singh [1*], Vinod Kumar [2], Rani Jain [3]

[1] Assistant Professor, Department of Agriculture, CCS Haryana Agricultural University, Hisar, India
[2] Junior Research, Department of Agriculture, CCS Haryana Agricultural University, Hisar, India
[3] Fellow Research, Department of Agriculture, CCS Haryana Agricultural University, Hisar, India

*Corresponding author: **Nidhi Singh**

Received: 12 Dec 2025
Revised: 28 Dec 2025
Accepted: 17 Jan 2026
Published: 08 Feb 2026

E-ISSN: 2989-5359
DOI: https://doi.org/10.54660/ejsa.2026.6.1.27-36

ABSTRACT

Background: The digitization of agriculture has resulted in the creation of large, varied datasets that include soil features, climate conditions, plant growth parameters and data about farm management. However, traditional ML approaches require these data to be compiled, raising serious questions regarding privacy, property rights, and legislation for individual farms and organizations.
Goal: The objective of this study is to review contemporary literature on Federated Learning (FL) as a decentralized concept that will allow for secure agricultural decision support systems without moving relevant information to a common database.
Framework: The literature review is based on peer-reviewed articles published mainly during the period from 2017 to 2026. and connects them to the larger context of precision agriculture driven by the Internet of Things and edge computing.
Findings: Based on the available literature, FML architectures find applications in crop yield forecasting, identifying diseases and pests in plants, irrigation planning, nitrogen management, and assessing climate risks, with techniques providing privacy protection such as differential privacy, secure aggregation, and trust protocols based on blockchain technology acting as supplementary measures. Comparative studies reviewed in this paper generally claim that FML has similar performance in prediction to centralized models while being able to reduce data exposure risk significantly and lessening communication costs in a number of cases. Among the challenges found in the literature are the problem of statistical heterogeneity (the issue of data being non-identically and non-independent), low connectivity in the countryside, difficulties with understanding aggregated models, and absence of unified standards governing the use of data in agriculture.
Significance: The review contributes to the existing body of knowledge by providing the audience with a structured synthesis of scientific literature on the issue of how FML can help achieve trustworthy digital agriculture.
Conclusion: Federated machine learning shows scientific promise that is reflected in extensive research on the issue.

Keywords: Decentralized learning; privacy-preserving artificial intelligence; smart farming; secure model aggregation; precision farming analytics; agricultural data governance; edge-IoT integration; distributed decision intelligence

1. Introduction

1.1. Digital Transformation of Agriculture

Agriculture has experienced an unbroken digital transition that has been made possible by sensor networks, remote-sensing systems, and powerful artificial intelligence (AI) applications. Data-based farming, or "smart" farming, utilizes data relating to space, time and image to aid in making decisions regarding irrigation, nutrient application, pest control, and harvest timing.

1.2. Agricultural Big Data and the Limits of Centralized Learning

Agriculture has experienced an unbroken digital transition that has been made possible by sensor networks, remote-sensing systems, and powerful artificial intelligence (AI) applications. Data-based farming, or "smart" farming, utilizes data relating to space, time and image to aid in making decisions regarding irrigation, nutrient application, pest control, and harvest timing.

1.3. Privacy Concerns in Agricultural Data Sharing

The agricultural data-sharing literature consistently identifies data ownership, privacy, and control as major barriers to collaborative data analytics, with concerns being most pronounced when R&D or commercial platforms depend on private ownership of sensor networks, proprietary germplasm information, or farmers' operational data. In places without advanced regulations, farmers and cooperatives have said they are unwilling to share raw data with third-party platforms over fears that integrated datasets might be used for commercial reasons or contain sensitive competitive information; thus, these issues have been seen as a continuing impediment to the widespread adoption of digital agricultural technologies among small and medium farms.

1.4. Federated Machine Learning as a Response

Federated Machine Learning (FML) has been formally presented as a way of collaborating among multiple clients - be it a mobile device, an edge server, or an institutional node - to train a single model using the guidance of a central server, while the training data of all clients remain locally with the respective owner clients. The original algorithm called Federated Averaging (FedAvg) has shown that rather than aggregating raw data, producer consumers can exchange model updates from the clients dispersed in space to build a good model, thus achieving significant gains in terms of time, communications, and privacy. Later researchers have made great progress in formalizing privacy guarantees, creating secure aggregation protocols, and developing ideas for dealing with statistical heterogeneity of client data (non-independent & non-identical distributions), thus making FML an established methodology used in the situations where data privacy, regulation, or bandwidth make it impossible to pool data in one place.

1.5. Secure Decision Support Systems in Agriculture

DSS for agriculture integrates sensor input, environmental factors, and management information for generating actionable advice on irrigation, fertilization, pest control, and harvesting. Studies on AI-oriented DSS indicate that farmers' confidence in such products depends on transparency, security measures, and data management, along with the system accuracy in prediction. Several papers indicate the reluctance of farmers to adopt such systems since they are seen as lacking transparency and undermining the control the farmers have over their data. The integration of privacy-preserving technologies into the decision-making process is viewed as a priority necessity for the next generation of agricultural AI.

1.6. Edge Computing and IoT Integration

Edge computing provides support for FML by means of local model development and inference on the site or in proximity to the agricultural IoT devices, bringing down the time taken, bandwidth needs, and the necessity of continuous connectivity. In applying the cloud-edge-device collaborative computing method in smart agriculture, reviews point out that the architecture uses fog or edge nodes as intermediate accumulators of information, which are responsible for the selection of devices participating in the process, coordinating updating of the models and transmitting information to the cloud coordination agencies. As a result, the multilayered architecture is relevant in the rural agricultural areas with discontinuous connectivity and the centralization of high-capacity information transfer.

1.7. Explainable and Trustworthy AI

In conjunction with privacy, the ability to explain AI-driven recommendations has become a determining factor affecting farmers' trust in and adoption of such innovations. Certain XAI techniques, e.g., model-agnostic explanation methods and attention visualization, have been invented that can help farmers and agricultural advisors understand how a certain recommendation was produced by the system and solve the problem of potential inconsistency with farmers' practical experience, which is generally a critical and widely recognized weakness of 'black-box' predictions. According to recent systematic reviews, the combination of XAI and FML is considered an emerging and promising approach to create agricultural AI systems that would be at the same time private, safe, and interpretable.

1.8. Global Research Status and Gaps

Numerous review literature studies have catalogued the way FML is being used in agricultural applications such as crop yield forecasting, plant disease classification, and nitrous oxide emission forecasting throughout North America, Europe, and Asia. There are a few common gaps that these reviews indicate: the lack of standardization of FML structures for agricultural applications, lack of consideration of the statistical heterogeneity due to different soil types and climate zones, little to no mention of explainability in combination with privacy-preserving aggregation, and a very limited number of papers discussing agricultural data governance and policies across various institutions.

1.9. Objectives of This Review

This article fills this gap in synthesis. In essence, it intends to: (i) outline the FML systems' architecture conceived in the context of agricultural decision support; (ii) provide an overview of the existing literature on FML application in various decision-making areas of agriculture; (iii) compile privacy-protecting techniques discussed in agriculture and related IoT/edge-computing literature; (iv) make a critical review of federated versus centralized learning paradigms showing various aspects highlighted in the selection of studies; and (v) identify the unsolved problems and perspective areas of research in the domain of secure, reliable, and explainable agriculture AI.

2. Review Methodology

This review synthesizes peer-reviewed literature identified through structured searches of Scopus-indexed and open-access scholarly databases, including PubMed Central, ScienceDirect, Springer Nature Link, IEEE Xplore, MDPI, and arXiv preprint repositories, using combinations of the keywords "federated learning," "agriculture," "precision farming," "privacy-preserving machine learning," "edge computing," "Internet of Things," "decision support systems," and "explainable AI." Publications were prioritized if they were peer-reviewed, published in English, and addressed federated or privacy-preserving learning architectures applicable to agricultural or closely related IoT/edge-computing contexts. Landmark methodological papers establishing the theoretical foundations of federated learning, differential privacy, and secure aggregation were included where they underpin concepts discussed in the agricultural literature, irrespective of publication venue. Reviewed studies were synthesized narratively and organized thematically around system architecture, data categories, application domains, and privacy mechanisms, consistent with established practice in interdisciplinary technology reviews; no primary experimental data were generated for this article, and no numerical performance results are reported as original findings. Table 1 summarizes the review's search and synthesis framework.

Table 1: Literature review search and synthesis framework

Component	Description
Databases searched	PubMed Central, ScienceDirect, Springer Nature Link, IEEE Xplore, MDPI, arXiv, Google Scholar
Search period	2016-2026, with emphasis on 2018-2026 publications
Core search terms	Federated learning; agriculture; precision farming; privacy-preserving AI; edge computing; IoT; decision support; explainable AI
Inclusion criteria	Peer-reviewed or archival preprint status; English language; relevance to agricultural or directly transferable IoT/edge FML architectures
Exclusion criteria	Studies lacking methodological detail on architecture, privacy mechanism, or application domain; non-English sources without translation
Synthesis approach	Thematic narrative synthesis organized around architecture, data categories, applications, and privacy mechanisms
Scope limitation	Conceptual and comparative synthesis; no primary experimental data, source code, or cryptographic implementation is presented

3. Federated Machine Learning Framework for Agricultural Decision Support

3.1. Categories of Agricultural Data Relevant to Federated Learning

The agricultural datasets described across the reviewed literature can be organized into several broad categories, each contributing distinct information to a federated decision-support pipeline. Soil property data (texture, organic matter, pH, nutrient status) inform fertility and irrigation-related recommendations. Weather and micrometeorological observations support risk assessment for disease pressure, frost, and drought. Remote sensing data, including satellite- and drone-derived vegetation indices, provide spatially continuous crop-condition information. IoT sensor streams capture near-real-time soil moisture, temperature, and equipment telemetry. Crop growth and yield records provide the historical outcome data against which predictive models are typically validated, while farm management logs describe input timing, cultivar choice, and operational practices. Table 2 summarizes these categories as conceptualized in the reviewed federated agricultural literature.

Table 2: Categories of agricultural data conceptually relevant to federated learning-based decision support

Data Category	Representative Variables	Primary Decision Relevance
Soil properties	Texture, pH, organic matter, macro/micronutrient status	Nitrogen recommendation, irrigation scheduling
Weather / micrometeorology	Temperature, precipitation, humidity, solar radiation	Disease risk, climate risk assessment
Remote sensing	Vegetation indices (e.g., NDVI-type indicators), canopy imagery	Yield prediction, pest and disease forecasting
IoT sensor streams	Soil moisture, in-field temperature, equipment telemetry	Irrigation scheduling, real-time monitoring
Crop growth / yield records	Phenological stage, historical yield outcomes	Yield prediction, model validation
Farm management logs	Input timing, cultivar selection, tillage practices	Farm management optimization

3.2. Federated Machine Learning Architecture

Conceptually, an FML system for agricultural decision support is composed of distributed client nodes - typically corresponding to individual farms, cooperatives, or regional aggregation points - each holding locally generated data that never leaves the client boundary. Each client trains a local model using its own data and transmits only the resulting model parameters, or a compressed representation of them, to a central or hierarchical aggregation server. The server combines these updates, commonly through variants of federated averaging, to produce an improved global model, which is then redistributed to participating clients for further local refinement. This iterative cycle continues across multiple communication rounds until the global model converges to an acceptable performance level. Reviews of federated learning in agriculture describe both horizontal federation, in which clients share the same feature space but different samples (for example, farms in different regions recording the same variables), and vertical federation, in which clients hold different feature sets describing overlapping entities, alongside hybrid and hierarchical variants that introduce intermediate fog or regional aggregation layers to reduce communication load. Figure 1 illustrates this conceptual workflow, and Figure 2 depicts a corresponding layered system architecture spanning farm-level data sources, edge computing, secure communication, cloud-based aggregation, and advisory interfaces.

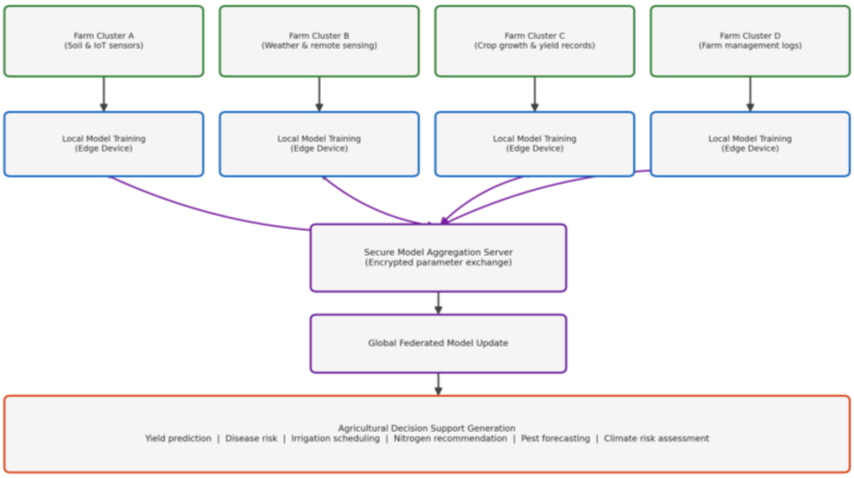

Fig 1: Conceptual workflow of a Federated Machine Learning framework for agricultural decision support, illustrating distributed farm data sources, local edge training, secure model aggregation, global model updating, and decision-support generation.

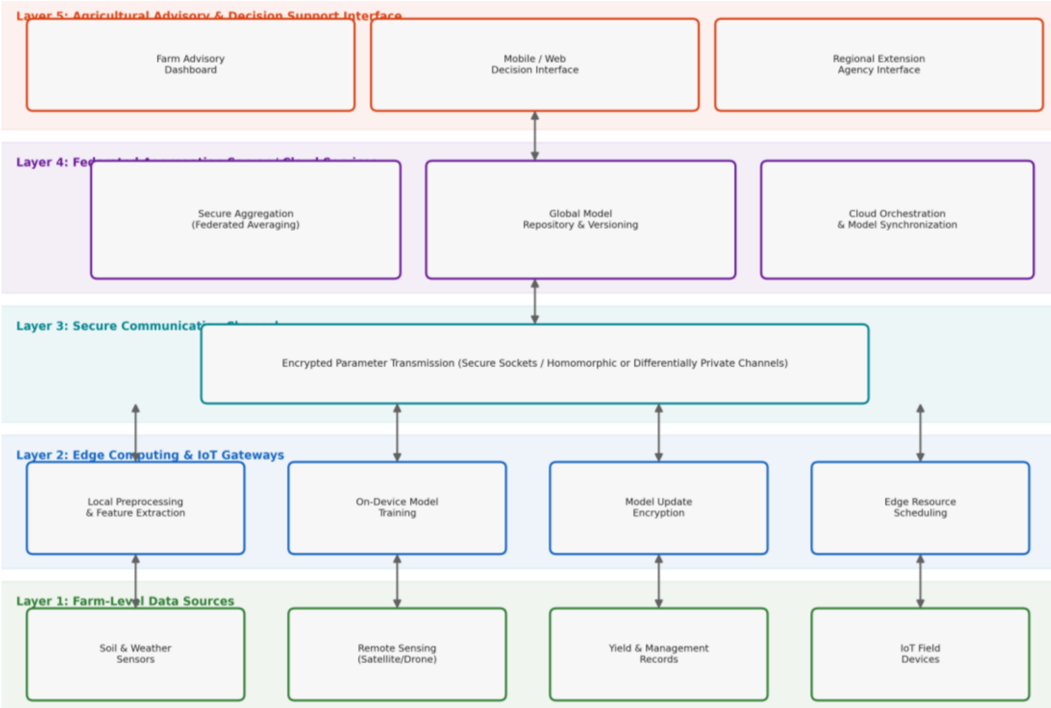

Fig 2: Layered system architecture showing the interaction among farm-level data sources, edge computing and IoT gateways, secure communication channels, the federated aggregation/cloud layer, and the agricultural advisory interface.

Table 3 summarizes the principal architectural components reported across the reviewed federated learning literature, together with their conceptual function within an agricultural decision-support pipeline.

Table 3: Core architectural components of Federated Machine Learning systems

Component	Conceptual Function	Illustrative Sources
Local client (farm/edge node)	Holds raw data locally; performs on-device model training	(McMahan *et al*., 2017) [1] (Žalik and Žalik, 2023) [3]
Aggregation server	Combines client model updates into an improved global model	(McMahan *et al*., 2017) [1] (Kairouz *et al*., 2021) [2]
Communication protocol	Governs the exchange of encrypted parameter updates between clients and server	(Kairouz *et al*., 2021) [2] (Konečný *et al*., 2016) [33]
Synchronization strategy	Determines round frequency and client participation (synchronous or asynchronous)	(Akbari *et al*., 2025) [10] (Abouaomar *et al*., 2025) [6]
Privacy-preserving mechanism	Adds formal or practical privacy guarantees to shared updates	(Bonawitz *et al*., 2017) [26] (Dwork and Roth, 2014) [27]
Edge/fog intermediary	Aggregates a subset of clients before forwarding to the cloud layer, reducing communication overhead	(Wang *et al*., 2025) [13] (Shen *et al*., 2024) [11]

3.3. Privacy-Preserving Communication and Security Mechanisms

The security of FML in agriculture rests on a combination of complementary mechanisms rather than any single technique. Secure aggregation protocols allow a server to compute the sum or average of client updates without observing any individual client's contribution, thereby preventing the server itself from reconstructing raw data from transmitted parameters. Differential privacy techniques introduce carefully calibrated statistical noise into shared updates, providing a formal, quantifiable guarantee that the presence or absence of any single data record cannot be confidently inferred from the aggregated output. Membership-inference and related attacks on machine learning models have demonstrated that, without such safeguards, adversaries can sometimes infer whether specific records were used in training, underscoring the practical necessity of these protections in agricultural contexts where farm-level data may be commercially sensitive. Blockchain-based trust and incentive mechanisms have also been proposed to support decentralized verification of client contributions and to discourage dishonest participation in federated agricultural networks, while fog-assisted architectures have been shown in the IoT literature to dynamically select aggregation nodes to balance security, latency, and resource constraints.

Table 4: Privacy-preserving mechanisms reported in federated and IoT/edge learning literature applicable to agriculture

Mechanism	Conceptual Description	Reported Benefit
Secure aggregation	Server computes aggregate model updates without accessing individual client contributions	Prevents server-side reconstruction of raw client data
Differential privacy	Calibrated statistical noise added to model updates or outputs	Formal, quantifiable bound on individual record inference risk
Homomorphic-style encrypted exchange	Model parameters exchanged in encrypted form across communication channels	Confidentiality of updates in transit
Blockchain-based trust/incentive protocols	Decentralized ledger for verifying and incentivizing honest client participation	Resistance to dishonest or malicious aggregation
Fog-assisted node selection	Dynamic selection of an intermediary aggregator among available fog nodes	Reduced latency and improved adaptability in constrained networks
Personalized/robust federated learning	Local model personalization layered on a shared global model to handle heterogeneous data	Improved resilience to non-IID farm-level data distributions

3.4. Decision Support Applications Enabled by Federated Learning

Across the reviewed literature, federated architectures have been described or proposed for a recurring set of agricultural decision domains. Crop yield prediction is the most extensively reviewed application, with federated frameworks applied to maize, wheat, rice, and soybean production using deep learning architectures such as convolutional and recurrent neural networks trained collaboratively across geographically distributed farms. Plant disease and pest detection has been addressed through federated computer-vision models trained on distributed image datasets, allowing detection models to benefit from diverse disease presentations across regions without centralizing potentially sensitive crop-health imagery. Irrigation scheduling and nitrogen management applications have been proposed using federated architectures that combine soil-moisture and weather data from multiple farms while respecting each farm's data locality. Climate risk assessment and nitrous oxide emission prediction have also been explored through IoT-integrated federated smart-farming architectures deployed on real sensor networks. Table 5 summarizes these application domains together with representative reviewed contributions.

Table 5: Decision support application domains reported for Federated Machine Learning in agriculture

Application Domain	Conceptual Approach Reported in Literature	Representative Sources
Crop yield prediction	Federated CNN/RNN/LSTM architectures trained collaboratively across multi-region farm clusters	(Hiremani *et al*., 2025) [5] (Abouaomar *et al*., 2025) [6] (Siniosoglou *et al*., 2023) [7]
Plant disease and pest detection	Federated computer-vision models trained on distributed, non-shared image datasets	(Alharbi and Csala, 2026) [18] (Pesantez-Cabrera *et al*., 2025) [8]
Irrigation scheduling	Federated models combining IoT soil-moisture and weather data across farms	(Akbari *et al*., 2025) [10] (Nguyen *et al*., 2023) [9]
Nitrogen / nutrient management	Federated regression-style models integrating soil and crop growth parameters	(Nguyen *et al*., 2023) [9] (Wang *et al*., 2025) [13]
Climate risk / emission assessment	IoT-integrated federated architectures for environmental prediction (e.g., nitrous oxide emissions)	(Akbari *et al*., 2025) [10]
Farm management optimization	Federated decision-support dashboards synthesizing multi-source recommendations	(Wolfert *et al*., 2017) [4] (Devi *et al*., 2025) [14]

4. Comparative Synthesis: Federated versus Centralized Learning Paradigms

Figure 3 presents a schematic comparison of the data and model flows characteristic of centralized and federated learning paradigms. In the centralized paradigm, raw data from each participating farm is physically transferred to and consolidated on a central server prior to model training, which the reviewed data-ownership and privacy literature identifies as a key source of farmer reluctance to participate. In the federated paradigm, only model parameters - and, where privacy mechanisms are applied, deliberately perturbed or encrypted versions of them - are exchanged, while raw data remains within each farm's own infrastructure throughout the process.

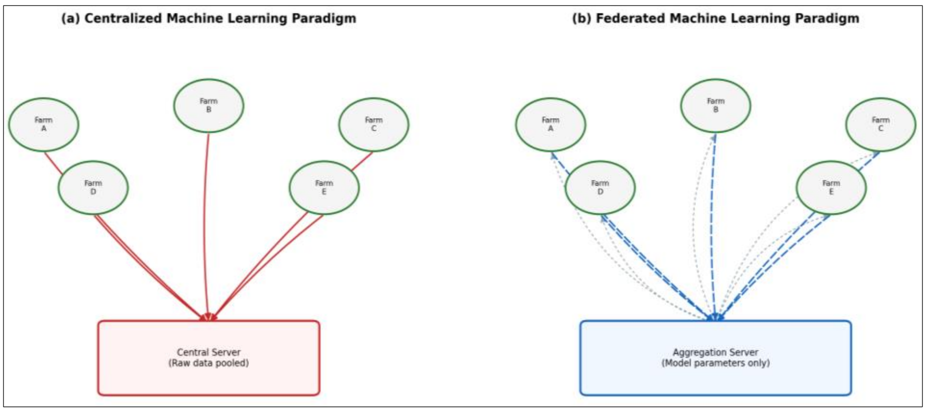

Fig 3: Schematic comparison of data and model flow in centralized versus federated machine learning paradigms, illustrating that federated architectures exchange only model parameters while raw farm data remains local.

The reviewed comparative and survey literature converges on several qualitative distinctions between the two paradigms, summarized in Table 6. Federated approaches are consistently reported as offering substantially improved data privacy and reduced raw-data exposure, since sensitive farm-level records never leave their originating device or institution. Several reviewed studies report that federated models can approach the predictive performance of centralized counterparts on comparable tasks, although this depends heavily on the degree of statistical heterogeneity across participating farms, with performance gaps widening under strongly non-independent and identically distributed (non-IID) data conditions. Communication efficiency is frequently cited as an advantage of federated architectures relative to naive centralized data transfer, particularly in bandwidth-constrained rural settings, though the iterative, multi-round nature of federated training introduces its own communication overhead that must be managed through techniques such as periodic averaging, update compression, and hierarchical fog-assisted aggregation. Scalability across a growing number of participating farms is generally reported as favorable for federated designs, though at the cost of increased architectural and orchestration complexity relative to a single centralized pipeline.

Table 6: Qualitative comparison of centralized and federated machine learning paradigms as characterized in the reviewed literature

Dimension	Centralized Learning (as reported)	Federated Learning (as reported)
Raw data exposure	Data physically pooled on a central server; higher exposure risk	Raw data remains local; only model updates are shared
Regulatory / ownership alignment	May conflict with data-ownership expectations of individual farms	Better aligned with data-locality and ownership preferences
Predictive performance	Often used as a reference benchmark in comparative studies	Frequently reported to approach centralized performance on comparable tasks
Sensitivity to data heterogeneity	Less affected, since all data is unified before training	Sensitive to non-IID data across farms; mitigation strategies are an active research area
Communication pattern	Single, often large, one-time (or batch) data transfer	Iterative, multi-round exchange of model parameters
Suitability for rural/low-bandwidth settings	Constrained by the need to transfer bulk raw data	Generally favored, especially with fog-assisted aggregation
Scalability across institutions	Coordination simplified but centralization risk increases with scale	Scales across many farms but adds orchestration complexity

5. Feature Relevance and Summary of Reviewed Studies

Beyond architecture and application, the reviewed literature offers a consistent, if largely qualitative, picture of which agricultural data categories are emphasized for particular decision domains. Figure 4 presents a conceptual relevance matrix synthesizing this pattern of emphasis across the reviewed studies; the matrix reflects the qualitative weight given to each data category within the cited literature rather than measured statistical correlations, and should be interpreted as an interpretive synthesis rather than an empirical result.

Fig 4: Conceptual relevance matrix summarizing the qualitative emphasis placed on agricultural data categories across reviewed decision-support domains, synthesized from the reviewed literature.

Weather and vegetation-index data are most consistently emphasized for disease-risk and pest-forecasting applications, reflecting the strong literature association between micrometeorological conditions, canopy condition, and pathogen or pest pressure. Soil properties and crop growth parameters are most frequently emphasized for yield prediction and nitrogen recommendation, consistent with agronomic understanding of nutrient-yield relationships. IoT sensor measurements are comparatively evenly weighted across domains, reflecting their role as a general-purpose, near-real-time data stream integrated into most federated agricultural pipelines described in the literature.

Table 7 synthesizes a representative cross-section of the studies reviewed for this article, summarizing their application domain, architectural approach, and reported privacy mechanism, without reproducing any numerical performance figures from the original sources.

Table 7: Comparative summary of representative reviewed studies on Federated Machine Learning in agriculture and adjacent domains

Study (First Author, Year)	Application Domain	FL Architecture / Approach	Reported Privacy Mechanism
(McMahan *et al*., 2017) [1]	Foundational FL algorithm (general)	Federated Averaging (FedAvg)	Local data retention; no raw data transfer
(Žalik and Žalik, 2023) [3]	Agriculture (general review)	Horizontal, vertical, and hybrid FL taxonomy	Architecture-dependent; reviewed comparatively
(Hiremani *et al*., 2025) [5]	Crop yield prediction	CNN/RNN/LSTM ensembles within FL	Local training with parameter-only sharing
(Abouaomar *et al*., 2025) [6]	Crop yield prediction (multi-crop)	Hierarchical, crop-clustered FL	Seasonal client clustering with local data retention
(Akbari *et al*., 2025) [10]	Nitrous oxide emission / smart farming	IoT-based federated smart-farming architecture	Privacy-aware sensor data aggregation
(Wang *et al*., 2025) [13]	Cross-domain smart agriculture	Fog-assisted privacy-preserving FL	Dynamic fog-node selection for aggregation
(Kotal *et al*., 2023) [21]	Agricultural data sharing policy	Policy-rule-enforced synthetic data generation	Rule-based confidentiality enforcement
(Bonawitz *et al*., 2017) [26]	General ML (foundational)	Secure aggregation protocol	Cryptographic secure aggregation

6. Discussion

6.1. Advantages of Federated Machine Learning for Agricultural Decision Support

The reviewed literature converges on privacy preservation as the principal advantage of FML relative to centralized alternatives, directly addressing the data-ownership concerns repeatedly identified as a barrier to collaborative agricultural analytics. By design, FML enables multiple farms, cooperatives, or research institutions to contribute to a shared, increasingly accurate model without any party gaining access to another's raw records, addressing a structural tension between the analytical benefits of large, diverse datasets and the legitimate proprietary interests of individual data holders described in the agricultural big-data literature.

6.2. Secure Agricultural Data Sharing and Distributed Intelligence

The convergence of FML with edge and fog computing further extends its relevance to agricultural settings characterized by intermittent connectivity, since local training and inference reduce dependence on constant, high-bandwidth links to distant servers. This distributed intelligence model also offers a degree of resilience: because no single node holds the complete dataset, the failure or compromise of any individual client does not expose the full federation's data, a property increasingly emphasized in reviews of IoT security in agriculture.

6.3. Comparison with Previous International Studies

Comparative reviews spanning studies conducted in the United States, India, China, Vietnam, Korea, and South American countries describe a broadly consistent trajectory: early federated agricultural applications focused narrowly on crop yield or disease prediction using relatively simple architectures, while more recent studies increasingly incorporate hierarchical aggregation, personalization for non-IID data, and explicit privacy mechanisms such as differential privacy and secure aggregation. This progression mirrors the broader federated learning literature, in which foundational algorithmic work has been progressively extended with formal privacy guarantees and robustness improvements for heterogeneous, real-world deployments.

6.4. Explainability and Trustworthiness

Despite these advances, the integration of explainable AI techniques with federated agricultural systems remains comparatively underdeveloped in the reviewed literature. Farmer trust research indicates that transparency and interpretability materially affect the adoption of AI-based recommendations, suggesting that future federated agricultural systems will need to pair privacy-preserving aggregation with accessible explanation mechanisms - for example, highlighting which data categories most influenced a given recommendation - rather than treating privacy and explainability as separate design concerns.

6.5. Practical Implementation Challenges

Several practical challenges recur across the reviewed studies. Statistical heterogeneity across farms - arising from differing soil types, climatic zones, crop varieties, and management practices - complicates model convergence and can bias global models toward better-represented regions unless mitigated through personalization or clustering strategies. Resource constraints on edge devices, including limited computational capacity and intermittent power or connectivity in rural areas, restrict the complexity of models that can be trained locally. The absence of unified data standards and interoperable platforms across agricultural technology providers further complicates federated collaboration across institutional boundaries. Finally, governance and incentive structures for federated participation remain underspecified in much of the literature, with blockchain-based incentive mechanisms proposed but not yet widely validated at agricultural scale.

6.6. Study Limitations

As a literature-based conceptual review, this article does not present new empirical results, and its synthesis is necessarily shaped by the scope and quality of the underlying reviewed studies, several of which report findings from simulated rather than large-scale operational deployments. The qualitative relevance matrix in Figure 4 reflects an interpretive synthesis of emphasis across the reviewed literature rather than a measured statistical outcome, and should be read accordingly. Because the federated learning and precision agriculture literatures are both expanding rapidly, this review reflects the state of published evidence at the time of writing and should be complemented by ongoing monitoring of the field.

6.7. Future Research Opportunities

Future research directions consistently identified across the reviewed literature include the development of standardized benchmarks for federated agricultural learning under realistic non-IID conditions; deeper integration of explainable AI methods within privacy-preserving aggregation pipelines; lightweight, resource-aware model architectures suited to low-power edge devices in rural settings; and the design of governance frameworks - potentially incorporating blockchain-based incentive and verification mechanisms - that formalize trust and participation among farms, cooperatives, and agricultural technology providers. Extending federated frameworks to incorporate multimodal data (combining imagery, sensor time series, and textual farm records) is also highlighted as a promising but comparatively early-stage direction.

7. Toward an Integrated Decision Support Framework

Synthesizing the architectural, application, and privacy-related themes discussed above, Figure 5 presents an integrated conceptual framework in which federated learning serves as the analytical core connecting geographically distributed farm regions to a shared set of decision-support outputs spanning irrigation scheduling, nutrient management, pest and disease prediction, yield forecasting, and climate-resilient farm management. Within this framework, each participating region retains full control over its underlying data, while the federated core continuously refines a shared global model that generates recommendations calibrated to the collective, privacy-preserved experience of all participating farms.

Fig 5: Integrated conceptual decision-support framework illustrating how federated learning connects geographically distributed farm regions to shared recommendations for irrigation, nutrient management, pest and disease prediction, yield forecasting, and climate-resilient farm management, while raw data remains local to each region.

Table 8 consolidates the challenges and corresponding mitigation directions discussed throughout this review, offering a structured reference for researchers and technology developers seeking to implement or extend federated agricultural decision-support systems.

Table 8: Persistent challenges in federated agricultural decision support and literature-identified mitigation directions

Challenge	Description	Literature-Identified Mitigation Direction
Statistical heterogeneity (non-IID data)	Differing soil, climate, and management conditions across farms bias global model convergence	Personalized federated learning; clustered or hierarchical aggregation
Edge-device resource constraints	Limited computation, storage, and power on rural IoT/edge devices	Lightweight model architectures; model compression and quantization
Connectivity limitations	Intermittent or low-bandwidth rural network access	Fog-assisted aggregation; asynchronous and periodic communication strategies
Data standardization gaps	Lack of interoperable formats across agricultural technology providers	Shared metadata standards; open agricultural data schemas
Governance and incentive design	Unclear frameworks for trust, participation, and accountability among federated partners	Blockchain-based incentive and verification protocols
Interpretability of aggregated models	Farmers may distrust recommendations from opaque global models	Integration of explainable AI methods within the federated pipeline

8. Conclusion

In this review, the current scientific literature regarding Federated Machine Learning has been summarized. The evidence presented shows that federated machine learning approaches consisting of distributed training, secure model aggregation, and privacy-preserving communications offer a sound scientific solution that allows for cooperating in model development among various farms and organizations with no need of sharing any potentially sensitive agricultural information. The fields of application of federated machine learning, from crop yield prediction to irrigation scheduling and climate impact mitigation, are numerous. collaboration related to institutional agricultural data. The paper outlines the significance of the areas that need to be further worked on, such as explainability for researchers, personalized learning in the environment filled with non-IID data, and development of low-weighted solutions. Emphasizing the importance of privacy by design principles for agricultural organizations and technology practitioners, as they influence farmers' trust and willingness to share their data, the paper brings the attention of policymakers in terms of the necessity to work on establishing interoperable standards and governance systems for agricultural federations embracing the idea of responsibility. With the development of digital agriculture processes, federated machine learning should find its proper place and help foster secure cooperation in the sphere of agricultural decision making as long as there are ongoing researches focusing on research issues related to discrepancies and governance mentioned in this review.

References

1. McMahan HB, Moore E, Ramage D, Hampson S, Arcas BA. Communication-efficient learning of deep networks from decentralized data. In: Proceedings of the 20th International Conference on Artificial Intelligence and Statistics. 2017;54:1273-1282.
2. Kairouz P, McMahan HB, Avent B, Bellet A, Bennis M, Bhagoji AN, *et al*. Advances and open problems in federated learning. Foundations and Trends in Machine Learning. 2021;14(1-2):1-210.
3. Žalik KR, Žalik M. A review of federated learning in agriculture. Sensors. 2023;23(23):9566.
4. Wolfert S, Ge L, Verdouw C, Bogaardt MJ. Big data in smart farming: A review. Agricultural Systems. 2017;153:69-80.
5. Hiremani V, Devadas RM, Preethi, Sapna R, Sowmya T, Gujjar P, *et al*. Federated learning for crop yield prediction: A comprehensive review of techniques and applications. MethodsX. 2025;14:103408.
6. Abouaomar A, El Hanjri M, Kobbane A, Laouiti A, Nafil K. Hierarchical federated learning for crop yield prediction in smart agricultural production systems. arXiv. 2025;2510.12727.
7. Siniosoglou I, Argyriou V, Bibi S, Lagkas T, Sarigiannidis P. Federated learning for crop trend forecasting across multiple farms. In: Proceedings of the IEEE International Conference on Big Data. 2023. p. 5321-5328.
8. Pesantez-Cabrera P, Kalyanapu A, Ganapathysubramanian B, editors. Enhancing smart farming through federated learning: Crop disease detection under heterogeneous environmental conditions. arXiv. 2025;2509.12363.
9. Nguyen TD, Nguyen LT, Pham QV. Federated learning in smart agriculture: An overview. In: Proceedings of the 15th International Conference on Knowledge and Systems Engineering (KSE); 2023 Oct; Vietnam. Piscataway (NJ): IEEE; 2023. p. 522-527.
10. Akbari M, Syed A, Kennedy WS, Erol-Kantarci M. IoT-based smart farming architecture using federated learning: A nitrous oxide emission prediction use case. ACM Journal on Computing and Sustainable Societies. 2025;3(2):1-24.
11. Shen Y, Guo H, Jan MA. Survey: Federated learning data security and privacy-preserving in edge-Internet of Things. Artificial Intelligence Review. 2024;57:154.
12. Basak S, Chatterjee K, Singh A, Kumar R. Privacy-preserving hierarchical fog federated learning for IoT intrusion detection. Frontiers of Computer Science. 2025;19:195903.
13. Wang H, Chen S, Zhao Y, Li M. Cloud-edge-device collaborative computing in smart agriculture: Architectures, applications, and future perspectives. Frontiers in Plant Science. 2025;16:1543210.
14. Devi DGP, Balachandra M, Kamath R. Explainable AI in agriculture: Review of applications, methodologies, and future directions. Machine Learning: Science and Technology. 2025;6(2):022001-022024.
15. Anonymous. Leveraging explainable AI for sustainable agriculture: A comprehensive review of recent advances. Artificial Intelligence Review. 2026;59:112.
16. Zambrano C, Thompson NM, Gunderson R, editors. Harnessing discrete choice experiments to elicit preferred configurations of trustworthy AI-augmented decision support systems for certified crop advisors. Frontiers in Artificial Intelligence. 2026;9:1627422.
17. Klerkx L, Rose D, Fielke S. Scale, trust, and the digital divide: A systematic review of AI and ML for agricultural applications. Frontiers in Artificial Intelligence. 2026;9:1594355.
18. Alharbi F, Csala D, editors. Integrating federated learning and explainable AI for plant disease detection: A systematic review of recent advances in precision agriculture. International Journal of Computational Intelligence Systems. 2026;19:145.
19. Gupta M, Abdelsalam M, Khorsandroo S, Mittal S. Security and privacy in smart farming: Challenges and opportunities. IEEE Access. 2020;8:34564-34584.
20. Elluri L, Piplai A, Kotal A, Joshi A, Finin T. A review of cybersecurity incidents in the food and agriculture sector. arXiv. 2024;2403.08036.
21. Kotal A, Piplai A, Elluri L, Joshi A. Privacy-preserving data sharing in agriculture: Enforcing policy rules for secure and confidential data synthesis. arXiv. 2023;2311.15460.
22. Torky M, Hassanein AE. Agriculture data sharing review: Opportunities, challenges, and future directions. Sustainability. 2025;17(3):1085-1109.
23. Kyveryga PM, Lowenberg-DeBoer J, Erickson B. Big data promises and obstacles: Agricultural data ownership and privacy. Agronomy Journal. 2022;114(5):2483-2492.
24. Ofori M, El-Gayar O. The role of big data in sustainable agriculture: Advancing environmental sustainability in precision farming systems. Cogent Food & Agriculture. 2026;12(1):2482686.
25. Carolan M. The politics of digital agricultural technologies: A preliminary review. Sociologia Ruralis. 2020;60(4):745-764.
26. Bonawitz K, Ivanov V, Kreuter B, Marcedone A, McMahan HB, Patel S, *et al*. Practical secure aggregation for privacy-preserving machine learning. In: Proceedings of the ACM SIGSAC Conference on Computer and Communications Security; 2017; Dallas, TX. New York: ACM; 2017. p. 1175-1191.

27. Dwork C, Roth A. The algorithmic foundations of differential privacy. Foundations and Trends in Theoretical Computer Science. 2014;9(3-4):211-407.
28. Shokri R, Stronati M, Song C, Shmatikov V. Membership inference attacks against machine learning models. In: Proceedings of the IEEE Symposium on Security and Privacy (SP); 2017; San Jose, CA. Piscataway (NJ): IEEE; 2017. p. 3-18.
29. Zhang C, Xie Y, Bai H, Yu B, Li W, Gao Y. A survey on federated learning. Knowledge-Based Systems. 2021;216:106775.
30. Zhao Y, Li M, Lai L, Suda N, Civin D, Chandra V. Federated learning with non-IID data. arXiv. 2018;1806.00582.
31. Dey T, Bera S, Paul B, De D, Mukherjee A, Buyya R. Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture. 2020;177:105709.
32. Boursianis AD, Papadopoulou MS, Diamantoulakis P, Liopa-Tsakalidi A, Barouchas P, Salahas G, *et al*. Internet of Things (IoT) and agricultural unmanned aerial vehicles (UAVs) in smart farming: A comprehensive review. Internet of Things. 2022;18:100187.
33. Konečný J, McMahan HB, Yu FX, Richtárik P, Suresh AT, Bacon D. Federated learning: Strategies for improving communication efficiency. arXiv. 2016;1610.05492.