

RESEARCH

Open Access



Federated agricultural yield forecasting under an Edge–Fog–Cloud framework: a multi-country privacy-preserving approach

Anil Utku^{1,2}, Abdulkadir Barut^{3,4*}, Hind Alofaysan⁵, Mohamed Djafar Henni⁶ and Azadeh Amoozegar²

*Correspondence:
Abdulkadir Barut
kadirbarut@harran.edu.tr

Full list of author information is
available at the end of the article

Abstract

Predicting agricultural yields is crucial for ensuring global food security and adaptation to climate change. However, the centralized collection of agricultural data across countries might pose governance, privacy, and data sovereignty issues. This research presents a privacy-preserving model for predicting agricultural yields using federated learning within an edge-fog-cloud integrated architecture. The experiments have been performed on a European cereal production dataset for the period 1990–2022, using a time-aware forecasting scheme. Centralized, edge-only, federated, and hybrid learning techniques have been systematically compared to examine the influence of distributed optimization on forecasting performance. It can be seen that central classical machine learning algorithms have yielded the best predictive performance. Federated deep learning models, meanwhile, showed relatively poor forecasting accuracy compared to their centralized counterparts under a non-IID distribution of multi-country agricultural data. Edge-only learning, meanwhile, has yielded comparatively low generalization capabilities due to the absence of collaborative learning.

Keywords Federated learning, Edge–Fog–Cloud computing, Agricultural yield forecasting, Non-IID learning, LSTM, Climate–yield modeling

1 Introduction

Agricultural yield prediction is an important aspect of food security, climate change, and economic stability [1]. Precise yield prediction allows the government and the Department of Agriculture to predict supply, trade, and climate change [2]. Using climate indices and agricultural production data, machine learning and deep learning algorithms have been widely adopted for yield prediction [3]. However, most current forecasting models use centralized data, meaning that agricultural and climate data from different regions are stored in a single database. While using centralized data for model training is beneficial, it also raises significant concerns about data sovereignty, data privacy, and cross-border regulatory issues [4]. In a multi-country environment, data collected from the agricultural production sector may be treated as private due to economic and



© The Author(s) 2026, modified publication 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

political factors. Therefore, unregulated data sharing across country borders may not be feasible.

On the other hand, agricultural data sources are highly heterogeneous. The occurrence of climate change, extreme weather events, agricultural practices, and data reporting standards is quite diverse across the globe [5]. Such differences cause non-identically distributed (non-IID) data patterns, making collaborative model training challenging. The conventional centralized model architecture ignores such governance and distribution patterns [6].

Edge and fog computing models have recently emerged as attractive alternatives to cloud-centric architectures. These models minimize latency, communication costs, and cloud dependence [7]. At the same time, federated learning (FL) has been explored as an innovative collaborative training approach for preserving data privacy [8]. FL enables multiple parties to cooperatively train a global model by exchanging model parameters instead of data [9]. However, there remain some research gaps in the growing body of literature on federated learning in agriculture. First, most existing studies focus on single-task learning, such as plant disease diagnosis, rather than large-scale yield prediction. Second, there is a lack of studies that compare centralized, edge-only, and federated learning in a unified manner. Third, the relationship between the complexity of the model architecture (classical machine learning, deep learning, and hybrid models) and the performance of federated learning remains inadequately investigated.

To fill the mentioned research gaps, this paper proposes an Edge-Fog-Cloud framework for multi-country agricultural yield forecasting. The proposed framework allows for centralized, edge-only, and federated learning in a unified manner. Based on a multi-decade European cereal yield dataset, we compare the performance of classical machine learning, deep learning, and hybrid models in the mentioned learning settings. The contributions of our work are threefold. First, we propose an Edge-Fog-Cloud architecture suitable for cross-country agricultural intelligence systems. Secondly, we develop and test a federated learning framework for multi-country yield prediction with strict temporal partitioning. Thirdly, we conduct a comparative study of classical, deep, and hybrid models to analyze the trade-offs between prediction accuracy, privacy protection, and stability of distributed optimization. Unlike previous FL-based agricultural forecasting studies, the proposed framework jointly evaluates centralized, edge-only, and federated learning paradigms under identical temporal forecasting conditions while integrating Edge-Fog-Cloud orchestration and hybrid temporal-structured feature learning. Our experiments show that although centralized ensemble learning methods achieve the best prediction accuracy, federated deep learning can approach centralized methods without sharing raw data. Our results offer valuable lessons for building privacy-preserving agricultural forecasting systems across multiple countries.

2 Related works

However, recent research tends to view smart agriculture through a distributed intelligence lens rather than a strictly cloud-focused approach. Yu et al. [10] in their work mention cloud-edge-device collaborative computing as a multi-layered paradigm, where the device/edge layer performs local inference and the cloud performs the heavy lifting. This clearly positions the paradigm within the context of addressing latency, bandwidth, and reliability challenges in agricultural applications. To provide a different perspective

on the above, a technical overview of smart agriculture enabled by edge computing, as presented by Gong et al. [11], focuses on multi-layered architectures that leverage the cloud's modeling capabilities and the edge layer's responsiveness and real-time processing requirements.

However, a significant body of literature considers federated learning a viable solution to address the limitations of data governance in agriculture, as participants may not be willing to share data on agricultural production, yields, and management. Dembani et al. [12] offer a general overview of data privacy in agriculture and federated learning, including the different forms of federated learning and their applications in agriculture, as well as the challenges associated with federated learning, which include data heterogeneity, communication complexity, and security. Similarly, Mohammed et al. [13] offer an overview of federated learning and smart agriculture, including the role of privacy and data heterogeneity.

From an architectural standpoint, Killeen et al. [14] present an IoT architecture for smart farming that leverages federated learning with edge/cloud computing for agricultural sensing and prediction applications, following the trend of distributed orchestration over cloud-centric FL. Yield prediction is arguably the most active field of FL research in agriculture. An extensive review of federated learning for crop yield prediction is offered by Hiremani et al. [15], which includes model families, data, metrics, and the main implementation challenges of FL for this application (communication overhead, non-IID data, interpretability).

In addition to these surveys, various research works have been conducted to propose end-to-end FL-based yield prediction system frameworks and compare them with their centralized counterparts. Dey et al. [16] proposed FLyer: A federated learning system for crop yield prediction with a focus on Agriculture 5.0, incorporating edge computing, FL-based parameter sharing, and engineering design to address latency/cost. Mukherjee et al. [17] conducted a comparative analysis of centralized and federated learning-based systems for crop yield prediction tasks using temporal neural networks such as LSTM/GRU, which is directly related to our comparison framework of centralized versus federated learning-based systems. Nirosha and Vennila [18] proposed a federated system for crop yield prediction using a combination of graph neural networks and recurrent neural networks, which represents a paradigm shift toward hybrid deep neural models for FL to effectively learn representations from distributed data.

Aside from algorithmic studies, some studies focus on FL design aspects in systems, particularly in the context of agriculture, which are critical to the overall results. In the work of Cowlshaw et al. [19], the authors investigate the effect of different levels of aggregation in FL (centralized and decentralized) and, using agricultural indicators derived from remote sensing images, demonstrate the significant impact of aggregation levels on overall results.

Recent studies have also explored federated and privacy-preserving intelligent systems in distributed cloud environments. AI-driven threat prediction models, malicious user prediction frameworks, and secure workload estimation architectures have been proposed to improve secure communication, adaptive orchestration, and distributed resource management in heterogeneous cloud infrastructures [20–24]. These studies further emphasize the growing importance of privacy-preserving collaborative intelligence and federated optimization strategies in distributed computing environments.

While yield prediction is one of the most popular FL prediction tasks, a substantial part of the agricultural FL literature focuses on plant disease diagnosis, which raises privacy concerns due to the use of image and operational data at the farm level. In the paper by Hari and Singh [25], the authors propose an adaptive federated deep learning method for diagnosing plant leaf disease using an intelligent weight transfer concept. The focus of the paper by Piccialli et al. [26] is the federated approach, AGRIFOLD, for the collaborative diagnosis and classification of plant leaf disease. The focus of the paper by Cris-tea et al. [27] is a hybrid federated transfer learning method for diagnosing tomato leaf disease, emphasizing its use without sharing the original data. The focus of the paper by Choubey et al. [28] is a lightweight federated transfer learning method for leaf disease diagnosis, emphasizing efficiency-driven design choices.

Recent federated learning studies have further explored privacy-preserving and heterogeneous learning scenarios in both medical imaging and smart agriculture applications. For instance, explainable federated deep learning frameworks integrated with Grad-CAM mechanisms have demonstrated strong performance for decentralized brain tumor classification under non-IID MRI data distributions [29]. Similarly, homomorphic encryption-enhanced federated learning architectures have been proposed for secure tomato leaf disease detection while protecting sensitive agricultural information against inference and reconstruction attacks [30]. Furthermore, hybrid federated few-shot learning frameworks combined with ensemble voting strategies have recently shown promising classification performance for cauliflower leaf disease detection under heterogeneous agricultural environments with limited labeled data [31].

These disease-related works validate two points that are generally applicable to the problem of yield prediction: (i) non-IID data is the norm, and (ii) there are benefits to be had with aggregation methods and light/hybrid modeling, not necessarily through FL itself.

From the literature, the key themes are edge/fog/cloud computing for reducing latency and improving operational viability, and federated learning for addressing data governance challenges. Nonetheless, several studies either (a) concentrate on a specific task type (notably disease detection) or (b) compare federated learning with a comprehensive assessment of centralized, edge-only, and federated learning approaches within a consistent forecasting framework. In light of these research challenges, our research makes several key contributions: (1) provides a formal Edge-Fog-Cloud process for agricultural forecasting, (2) compares centralized, edge-only, and federated learning approaches under strict temporal partitioning, and (3) compares machine learning, deep learning, and hybrid models on a multi-country European cereal dataset.

However, despite promising results from some federated learning approaches already applied in agriculture and healthcare, several limitations remain worth considering. The vast majority of related studies focus on image and disease classification rather than on predicting agricultural yields over time under heterogeneous, multi-country conditions. Moreover, many current studies operate within the framework of centralized optimization or homogeneous client distributions, which are unrealistic in practice when addressing the issue of decentralized agriculture. In addition, there is a lack of investigation into federated learning techniques for privacy preservation, temporal forecasting, and heterogeneity in client distribution.

Despite the promising results from previous federated learning research in agriculture and other distributed artificial intelligence applications, the methodologies, privacy, and heterogeneity of non-IID environments vary significantly across studies. Moreover, previous studies generally focus on image classification rather than long-term forecasting of agricultural yields. To contextualize the contribution of the proposed study within recent federated learning research, a comparative analysis of feature similarities between the previous and proposed studies is presented in Table 1.

As shown in Table 1, most previous federated learning studies focus on classification tasks under partially decentralized approaches. In contrast, the proposed study aims to address the problem of agricultural yield forecasting across multiple countries within a privacy-preserving framework in a hybrid Edge-Fog-Cloud system. Moreover, the proposed study provides a comprehensive comparison of centralized, edge-only, and federated learning techniques, as well as classical machine learning, deep learning, and hybrid model approaches, within a distributed optimization framework.

In this regard, it is important to note that the federated learning approaches for the sequence-based DL-LSTM architecture were primarily used in the present study due to their applicability to gradient-based optimization via FedAvg. Ensemble machine learning approaches, including Random Forest and XGBoost algorithms, were used as strong centralized baselines for tabular forecast tasks; nevertheless, no federated approaches for tree-based ensemble learning were tested in the present simulations. Thus, the emphasis was on the effect of federated distributed optimization on deep sequence forecasting models.

Despite many recent federated learning research efforts, for example, FLYer, Mukherjee and Buyya, and Cowlishaw et al., there are still many key research gaps related to the potential applicability of distributed learning in the context of agriculture. Current research work mostly tends to concentrate on federated crop prediction or distributed agricultural sensing independently, while not comparing various learning approaches, including federated and conventional, based on the same time-based problem setting. Another common drawback is that most research studies do not provide a detailed analysis of model complexity across various model types, ranging from classic machine learning to deep and hybrid architectures, in heterogeneous, multi-country, non-IID

Table 1 Comparative analysis of related federated learning studies

Study	Application domain	Learning paradigm	Privacy-preserving	Non-IID analysis	Model type	Main limitation
Brain tumor FL study	Medical imaging	Federated	Yes	Partial	CNN	No temporal forecasting
HEFL tomato study	Agriculture	Federated	Yes	Limited	CNN	Image classification only
Fed-FSL study	Agriculture	Federated	Partial	Yes	Hybrid ensemble	Limited temporal analysis
Proposed study	Agricultural yield forecasting	Federated Edge-Fog-Cloud	Yes	Yes	Hybrid DL-LSTM	Moderate FL performance degradation

settings. Unlike previous work, our study adopts an integrated approach to federated forecasting across edge, fog, and cloud.

For further clarification of the differences between the proposed framework's methodology and past federated agricultural learning research studies, a comparative analysis is presented in Table 2.

As depicted in Table 2, past federated agricultural learning research has primarily focused on federated prediction settings or distributed sensing tasks. However, the proposed framework considers all three learning methods (centralized, edge-only, and federated) within a single unified Edge-Fog-Cloud paradigm, while also comparing classical machine learning, deep learning, and hybrid approaches.

3 Material and method

This section introduces the materials, architectural design, and methodological framework used in this study to design and evaluate a privacy-preserving agricultural yield forecasting system. The proposed system combines multi-country cereal yield data and climate indices in a distributed Edge-Fog-Cloud computing framework. A federated learning framework is developed to facilitate collaborative model training without the need to share raw data across country borders. To provide a comprehensive comparison of the proposed system, centralized, edge-only, and federated learning frameworks are systematically compared using the same temporal evaluation framework. The methodology framework comprises data preprocessing, sequence formation based on a fixed multi-year time window, deep learning-based sequence modeling, federated aggregation strategies, and multi-metric system performance evaluation. The aim of the methodology used is to evaluate whether distributed learning could achieve the same performance as a central system in the context of multi-country agricultural yield forecasting and data sovereignty.

3.1 Problem definition

The problem this research addresses is multi-country agricultural yield forecasting, subject to data sovereignty and distributional heterogeneity constraints. The problem

Table 2 Comparative analysis of existing federated agricultural forecasting frameworks and the proposed study

Study	Dataset	Task	FL architecture	Edge/Fog/Cloud support	Model type	Privacy mechanism	Evaluation setting
[16]	Agricultural IoT datasets	Crop prediction	Federated edge learning	Partial	DL	Parameter sharing	Distributed simulation
[17]	Crop yield datasets	Yield prediction	Federated learning	No explicit fog layer	LSTM/GRU	Parameter aggregation	Centralized vs FL
[19]	Remote sensing agricultural data	Agricultural monitoring	Hierarchical FL	Partial	ML/DL	Aggregated updates	Multi-level aggregation
Proposed study	European cereal dataset	Multi-country yield forecasting	Edge-Fog-Cloud FL	Full	Classical ML + DL-LSTM + Hybrid	Privacy-preserving parameter sharing	Centralized vs edge-only vs federated under strict temporal split

is formulated as a distributed time-series regression problem, where each participating country has its own historical agricultural and climate data. The data have a panel structure, with a spatial component representing countries and a temporal component spanning several decades. The response variable is annual agricultural yield, predicted using a fixed-size sliding window of climate and agricultural features. In this research, a five-year temporal context is used to account for the lagged effects of climate and temporal dependencies in agricultural production dynamics.

In contrast to typical centralized learning scenarios, the raw data in this research cannot be easily aggregated across countries due to regulatory, organizational, and governance factors. This implies that the learning problem should be reformulated as a distributed optimization problem, where each country trains a local model on its own data while jointly constructing a global model. Only model parameters or updates are communicated, while raw agricultural data are retained locally. This turns the agricultural yield forecasting problem into a cross-silo federated learning problem, where each country acts as a stable, sufficiently large client.

One of the defining features of this problem is that the data is not identically distributed. Crop yields, climate variability, management practices, and extreme weather patterns vary widely across countries. This results in variations in yield scales, statistical distributions, noise levels, and climate-yield response functions. As such, the learning framework must accommodate these two phenomena while ensuring stable convergence during collaborative optimization.

In order to evaluate the effectiveness of distributed learning under these circumstances, the forecasting problem is evaluated using three different strategies, namely, centralized learning, where all data from all countries is used to form one single dataset; edge learning, where each country uses its own data to train its own model; and federated learning, where collaborative learning occurs without access to raw data. The research question is whether federated learning can achieve forecasting accuracy comparable to that of centralized learning.

The target, therefore, is not merely to minimize the prediction errors. Rather, the problem at hand is to achieve a trade-off among accuracy, privacy preservation, scalability, communication costs, and robustness to non-IID data distributions. The proposed framework is expected to optimize these trade-offs within a real-world Edge-Fog-Cloud computing paradigm for multi-country agricultural forecasting scenarios.

3.2 Dataset

In this research, a panel dataset is used that combines agricultural production data from the World Bank's Climate Change Information Portal (WB-CCKP) along with climate indicators [32]. The dataset spans 1990 to 2022 and covers some European countries and three important cereal crops: barley, wheat, and maize. The dataset has agricultural indicators such as yield, harvest area, production quantity, climate indicators such as temperature, precipitation, drought, extreme events, geographic indicators such as country-level observations, and time indicators such as annual observations. The independent variables comprise agricultural variables (YIELD, AREA, PRODUCTION), crop variables, country variables, and several climatic variables derived from the WB-CCKP data set, including precipitation, temperature, drought, heat stress, and extreme climate variables. The target prediction variable corresponds to annual crop yield measured in

kilograms per hectare (kg/ha). The final dataset contains approximately 3,200 observations and over 40 explanatory variables.

The federated learning experiment used 39 country-level clients from Europe, which include Albania, Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, Türkiye, Ukraine, and other European countries producing cereals that fulfill minimal temporal data requirements.

The total size of the dataset in terms of the number of rows (observations) and columns (features) is shown in Fig. 1.

As shown in Fig. 1, the proposed system employs a hierarchical Edge-Fog-Cloud architecture that supports privacy-preserving distributed agricultural forecasting. In such a model, each country participates in the forecast process as an independent edge client and trains its own models using its agricultural and weather data. No data is uploaded; only the model parameters are uploaded to the Fog layer. The Fog layer uses the FedAvg approach for federated averaging and creation of the global model. The updated model is downloaded back to each client for further training cycles. The distribution of records across crop types is shown in Fig. 2.

Figure 2 shows that the number of records for Barley and Wheat is similar, while Maize has a few fewer. The geographical distribution of the data across the countries is shown in Fig. 3.

The dataset shows a good geographical distribution across European countries. It is important to understand the distribution of the target variable (YIELD) before modeling. The distribution of the target variable is shown in Fig. 4 as a histogram.

The distribution has moderate right skewness, with most values between 2,000 and 6,000 units, while the highest values exceed 10,000 units. The distribution is not normal, which warrants the use of non-linear models such as ensemble learning and deep neural networks. Figure 5 below shows the mean annual yields for each crop type.

All three crops have a general increasing trend. This is an indication of technological progress, better agricultural practices, and adjustment to climatic conditions. Figure 6 shows the mean yield over time for the top six countries with the most observations.

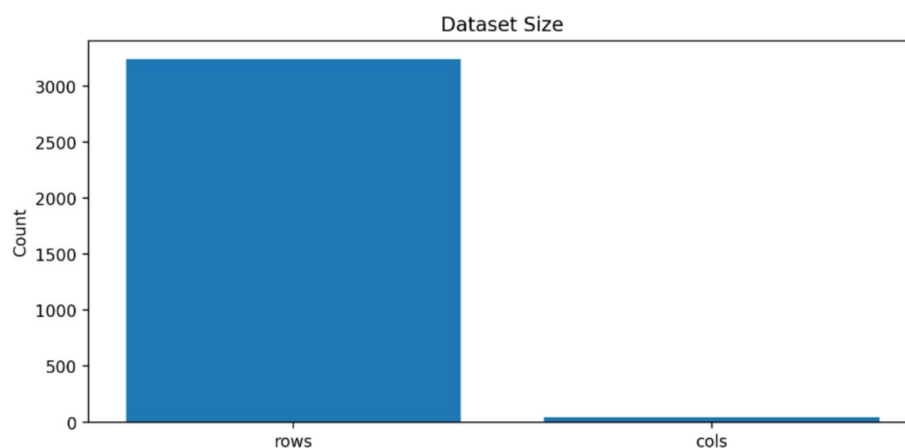


Fig. 1 Overview of dataset dimensions (rows and columns)

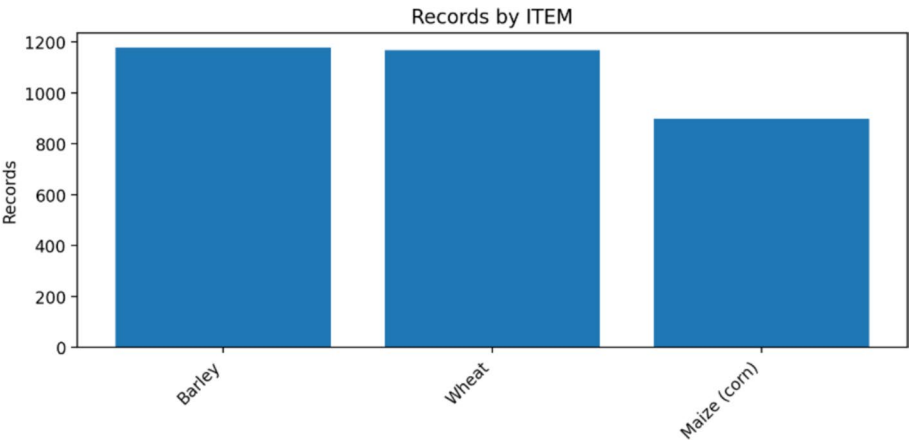


Fig. 2 Number of records by crop type (Barley, Wheat, Maize)

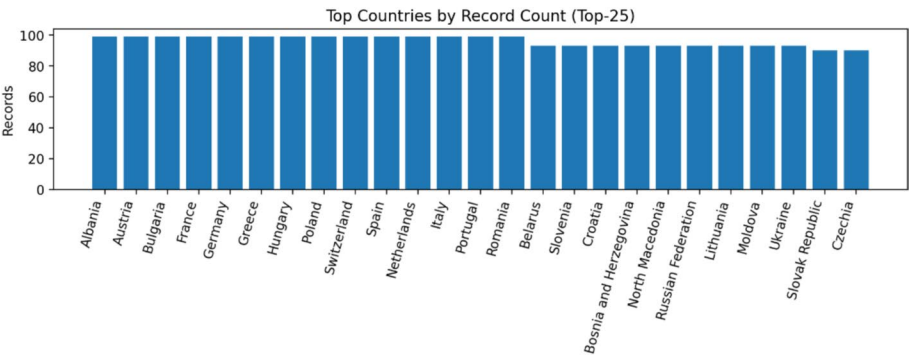


Fig. 3 Distribution of records across the top 25 participating countries

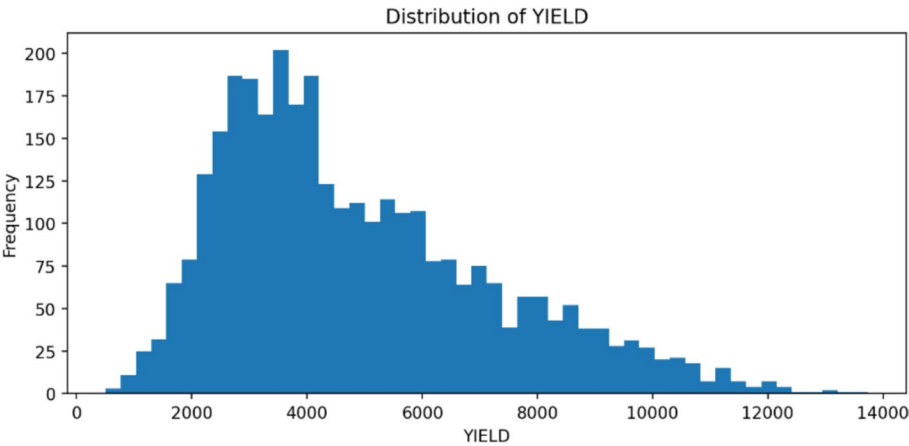


Fig. 4 Distribution of annual crop yield values

The figure above shows variability among countries, which can be attributed to differences in responses to climatic conditions and agricultural management practices. Figure 7 shows the evolution of mean climate indices over time.

Some temperature and extreme event variables have shown an increasing trend over time, likely due to increased climate variability during the study period. To analyze the

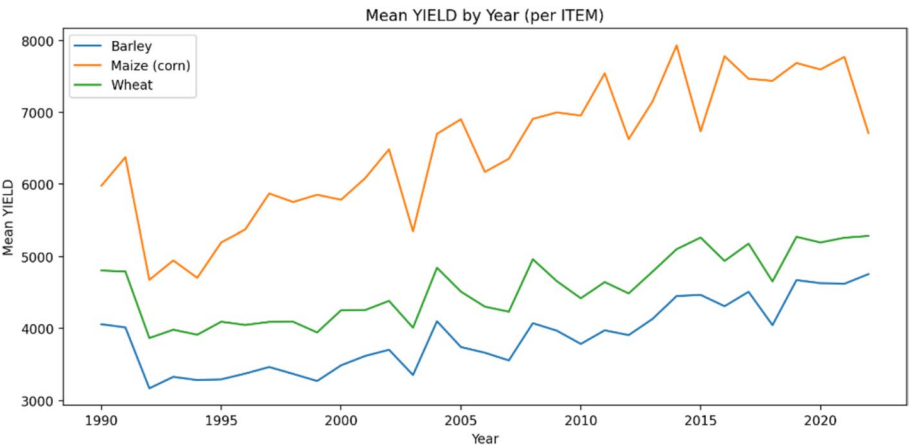


Fig. 5 Mean YIELD by year for Barley, Wheat, and Maize

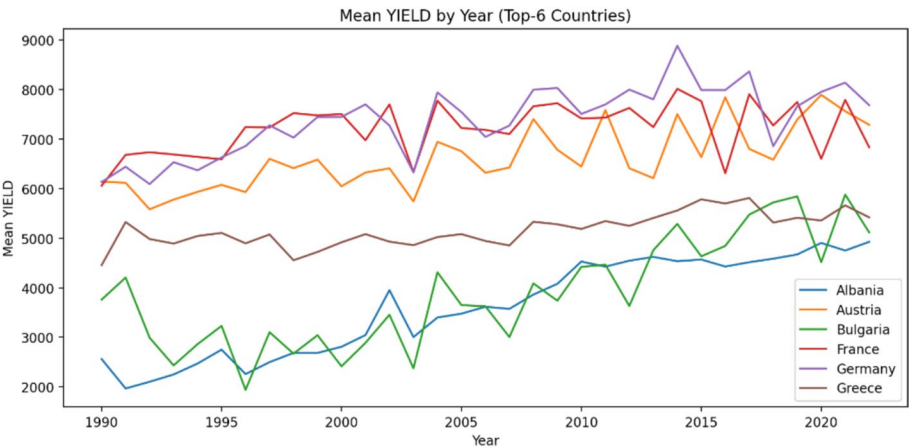


Fig. 6 Mean YIELD by year for the top-6 countries

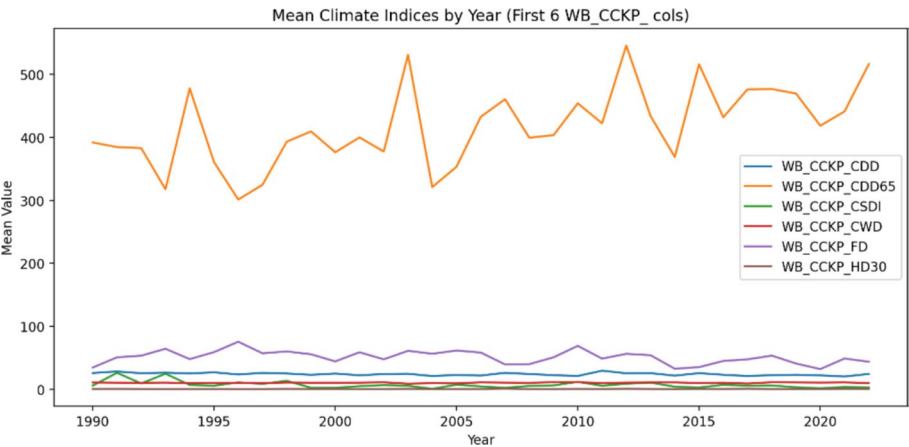


Fig. 7 Temporal evolution of selected climate indices

correlation between agricultural production and climate variables, a Pearson correlation matrix was calculated.

From the heatmap, there are moderate correlations between yield and temperature variables, high correlations between precipitation variables, and signs of multicollinearity among various climate variables. The results provide a rationale for the application of sophisticated machine learning and deep learning algorithms that can effectively model high-dimensional correlated feature spaces (see Fig. 8).

Interpolation techniques were used to fill in missing data within the climatic and agricultural indices based on temporal continuity and variable availability. For climatic variables that exhibited temporal continuity, linear interpolation was applied for short temporal gaps, while the remaining missing values were filled with the median statistics of the respective training partitions.

Variables that could not be accessed prior to the prediction year were excluded from the prediction process to avoid data leakage. Historical agricultural data, such as production volume and harvested area, were used only for past years within the five-year input period.

To facilitate reproducibility, a more thorough description of the agricultural forecasting data utilized for this experiment is provided in Table 3. This includes information on the time span, participating countries, crop categories, feature groups, preprocessing techniques, and other parameters of our federated learning framework.

According to the description in Table 3, our federated dataset consists of heterogeneous, multi-country agricultural and climate data collected over several decades, under non-IID assumptions. The incorporation of both climate-based temporal features and

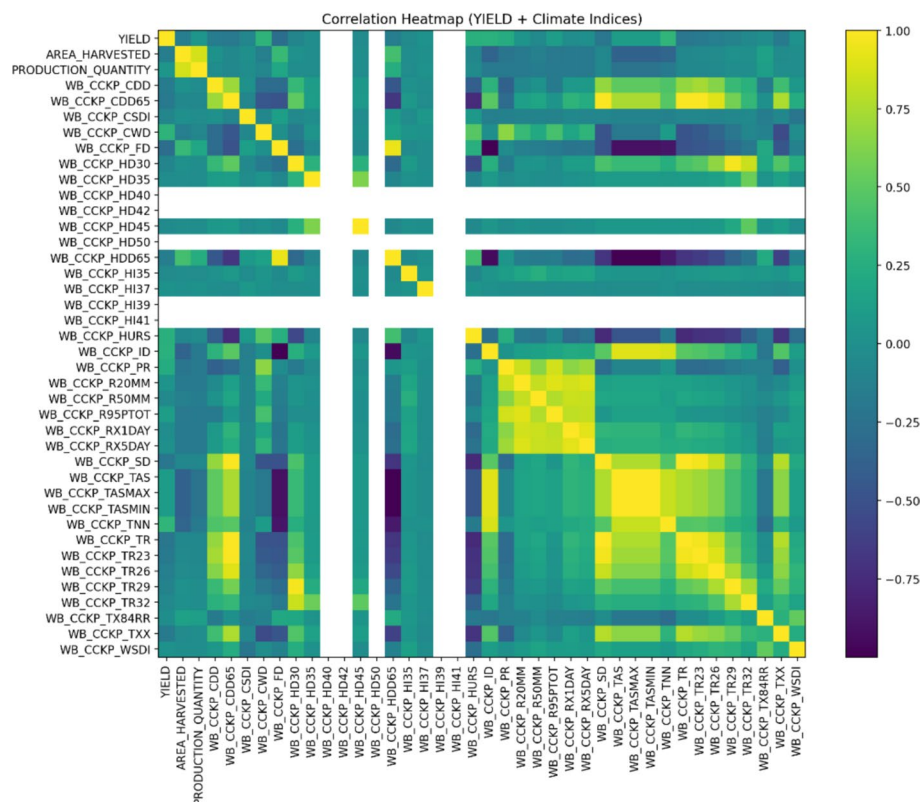


Fig. 8 Pearson correlation matrix of yield and climate variables

Table 3 Summary of the agricultural forecasting dataset used in this study

Attribute	Description
Temporal coverage	1990–2022
Number of country-level clients	39
Crop categories	Barley, Wheat, Maize
Prediction target	Annual crop yield
Target unit	kg/ha
Total observations	~ 3200
Number of explanatory variables	> 40
Climate variables	Temperature, precipitation, drought, heat stress, extreme events
Agricultural variables	Yield, harvest area, production quantity
Data granularity	Annual country-level observations
Missing-value handling	Interpolation + median imputation
Normalization	Z-score standardization
Evaluation protocol	Strict time-aware split
Training period	1990–2014
Validation period	2015–2018
Testing period	2019–2022
Federated partitioning	Country-based decentralized partitioning

agricultural attributes facilitates the realistic evaluation of central, edge-based, and federated forecasting algorithms. In addition, proper leakage-prevention methods and time-based splits were used to conduct realistic, reproducible forecasting.

3.3 Proposed Edge–Fog–Cloud architecture

This section describes the proposed Edge-Fog-Cloud architecture for privacy-preserving, scalable, and latency-conscious agricultural yield forecasting across heterogeneous country-level data distributions. The proposed architecture is capable of supporting three different operational modes: centralized learning, edge-only learning, and federated learning, and is mainly designed to be optimal for federated learning, where the raw data is kept local, and a global model is learned collaboratively.

The proposed Edge-Fog-Cloud architecture is depicted in Fig. 9. The proposed architecture is intended to support privacy-preserving collaborative learning across multiple sovereign states. The architecture is also intended to ensure scalability, efficient communication, and global lifecycle management of the model. It is worth noting that the suggested Edge-Fog-Cloud learning system serves as a conceptual and simulated architecture intended to emulate decentralized agricultural federated learning systems rather than a physically implemented real-world infrastructure.

As shown in Fig. 9, the proposed architecture comprises three layers. The Edge layer supports local data processing and model training in each country. The Edge layer ensures that raw agricultural and climate data are processed locally in each country. The Fog layer supports federated learning rounds and model update aggregation using Federated Averaging (FedAvg). The Cloud layer supports global services such as model registration, experiment tracking and monitoring, and large-scale simulation.

3.3.1 Design objectives and rationale

Multi-country agricultural forecasting is challenged by three major issues: (i) the massive scale of spatiotemporal data, (ii) the strict constraint on data sharing across countries, and (iii) non-IID (non-identically distributed) data due to varying climate patterns,

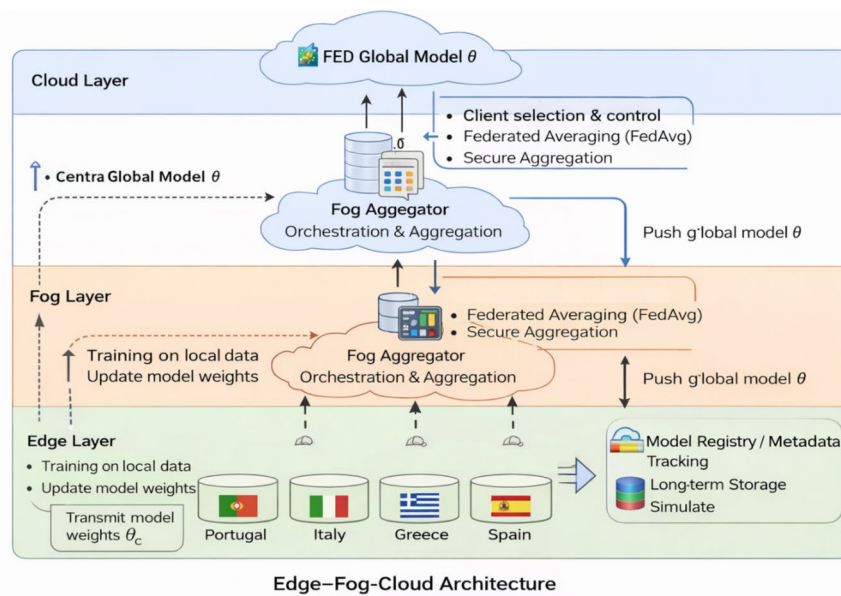


Fig. 9 The proposed Edge-Fog-Cloud architecture

agricultural practices, and reporting mechanisms. The centralized learning paradigm is convenient for learning but requires transferring raw data to a centralized data center, which is not privacy-friendly, scalable, or bandwidth-efficient. On the other hand, edge-only learning is privacy-friendly but often suffers from data scarcity and limited generalization.

The architecture allows for learning of models via unified optimization and retains sensitive data on-device/local servers (Edge). The architecture reduces cloud dependency and communication costs by introducing a Fog layer for data orchestration, collection, and analysis close to the source. The Cloud layer is used to store model artifacts, track experiments, and run computationally intensive tasks such as hyperparameter tuning, large-scale testing, and digital twin simulation.

3.3.2 Layered architecture overview

The system is divided into three tiers. **Edge Layer (Country Nodes):** Local training and inference on each country's dataset. **Fog Layer (Regional Aggregators/Orchestrators):** Training rounds, secure aggregation, and optional hierarchical federation. **Cloud Layer (Global Services):** Model registry, monitoring, metadata catalog, and offline benchmarking. Each tier is explained below.

3.3.3 Edge layer: country-level learning nodes

Each Edge node represents a sovereign entity (country, ministry, or agricultural agency) that maintains historical agricultural and climate data. The edge node performs the following tasks.

Local Data Management and Feature Construction: The edge nodes process annual data (1990–2022 in our work) to construct supervised learning examples based on a fixed context window size. For a window size of 5 years, the training example is constructed as follows:

Input sequence:

$$X_{t-5:t-1} \in \mathbb{R}^{5 \times d} \quad (1)$$

Target:

$$y_t \in \mathbb{R}(\text{yieldatyear } t) \quad (2)$$

where c denotes the country-level federated client, t represents the target prediction year, X_t corresponds to the feature vector at year t , d is the feature dimension, and $X_{t-5:t-1}$ denotes the five-year temporal input sequence used for forecasting.

This transformation yields a country-specific training set:

$$D_c = \{(X_i, y_i)\}_{i=1}^{N_c} \quad (3)$$

Here, N_c denotes the total number of supervised training samples available for country-level client c . To maintain numerical stability during neural training, the edge nodes perform a feature normalization technique (z-score scaling). In the federated setup, normalization parameters can be calculated on the training set and communicated as global values or calculated locally with proper reporting in the threat model (see Section on privacy assumptions).

Local Model Training:

At each federated round r , the edge node receives the current global model parameters $\theta^{(r)}$ and performs local optimization for E epochs:

$$\theta_c^{(r+1)} = \theta^{(r)} - \eta \sum_{e=1}^E \nabla L(\theta; B_{c,e}) \quad (4)$$

where $\theta^{(r)}$ denotes the global model parameters at communication round r , $\theta_c^{(r+1)}$ represents the locally updated model parameters of client c , $B_{c,e}$ denotes the mini-batches sampled from country-level client c during local epoch e , and η is the learning rate. In this research, the deep model is sequence-based (LSTM) to incorporate temporal dependencies; further, a discrete embedding can represent crop type to facilitate parameter sharing for different crops while maintaining their respective dynamics.

Privacy-Preserving Update Transmission: Rather than transmitting the actual data, each edge node will transmit one of the following to the Fog layer. Model weights: $\theta_c^{(r+1)}$ or model deltas/gradients:

$$\Delta \theta_c^{(r+1)} = \theta_c^{(r+1)} - \theta^{(r)} \quad (5)$$

where $\Delta \theta_c^{(r+1)}$ represents the parameter update transmitted from client c to the Fog aggregation layer. Only parameter updates are transmitted, which reduces data exposure and satisfies real-world governance requirements.

3.3.4 Fog layer: orchestration and aggregation

Fog Layer (Orchestration and Aggregation): The Fog layer is the control plane and aggregation layer, hosting functionalities that are expensive or inefficient to execute at the edge, but not required to be pushed all the way to the cloud. It supports two deployment modes. Single Fog aggregator: one orchestrator for all edge nodes. Hierarchical federation: several regional Fog aggregators aggregate regionally, followed by global aggregation in the cloud (optional extension). The Fog orchestrator triggers rounds

$r = 1, 2, \dots, R$, selects a subset of clients $S_r \subseteq \{1, \dots, C\}$ based on availability, minimum sample thresholds, and potential balancing constraints:

$$|S_r| = \max(1, \lfloor C \cdot \rho \rfloor) \quad (6)$$

where C denotes the set of all participating clients, S_r represents the subset of selected clients at communication round r , and ρ is the client participation ratio. This step is essential under realistic edge constraints (intermittent connectivity, limited compute windows).

Weighted Model Aggregation (FedAvg):

The Fog layer calculates the global model using the Federated Averaging (FedAvg) optimization strategy proposed by McMahan et al. [33] after collecting parameter updates from selected clients:

$$\theta^{(r+1)} = \sum_{c \in S_r} \frac{N_c}{\sum_{k \in S_r} N_k} \theta_c^{(r+1)} \quad (7)$$

Here, n_c denotes the number of local samples owned by client c , $\theta_c^{(r+1)}$ represents the locally updated client model parameters, and $\theta^{(r+1)}$ corresponds to the aggregated global model parameters after federated averaging. This weighted average by sample count ensures that clients with more data samples contribute more to the global update.

Robustness and stability controls: Convergence might become unstable with Non-IID data distributions. The Fog layer can leverage stability controls like learning rate control and early stopping triggers based on validation telemetry, Huber loss (robust to outliers) for regression tasks with heavy-tailed targets, client selection (minimum sample thresholds, update vector bounds), and aggregation clipping (optional), and the update magnitude to stabilize the global model. These controls prevent the risk of divergence due to outlier clients or abrupt regime changes [34].

Optional secure aggregation: For enhanced privacy, the Fog layer can use secure aggregation protocols such that client updates are never exposed in plaintext form to the aggregator [35]. Although the present experimental setup is centered on functional FedAvg, the design is amenable to secure aggregation modules (such as additive masking approaches).

3.3.5 Cloud layer: global services and lifecycle management

The Cloud layer conceptually provides persistent services that support lifecycle management, monitoring, simulation, and reproducible experimentation within the proposed distributed learning framework [36]:

Model registry and versioning: All global models $\theta^{(r)}$ are versioned with metadata that includes training parameters, client participation, and validation scores [37]. This facilitates reproducible experiments, rollback to stable checkpoints, and strategic deployment of model versions to edge nodes.

Monitoring, drift detection, and benchmarking: The cloud provides monitoring dashboards for validation RMSE/ R^2 over rounds, country-level performance summaries, and data drift indicators (e.g., distribution shifts in climate indices). Moreover, the cloud performs offline benchmarking against centralized baselines and traditional ML models.

Digital twin/simulation framework: For systematic analysis prior to deployment, the cloud provides a simulation framework that can model client availability patterns, communication delays, compute constraints, regional clustering strategies (hierarchical fog), and non-IID intensity across countries [38]. This is an important part of analyzing the trade-offs between accuracy, communication cost, privacy, and latency. In the current study, these functionalities were evaluated through simulation-based federated orchestration rather than real-world physical infrastructure deployment.

3.3.6 End-to-end workflow

The end-to-end training/inference process follows these steps: Initialization (Cloud → Fog): The global model is initialized and shared across the cloud and fog. Round orchestration (Fog): For each round, a set of clients is chosen. Local training (Edge): Each chosen country trains for E epochs on local data using window = 5 sequences and updates the weights. Aggregation (Fog): Updates are aggregated using FedAvg to compute a new global model. Validation and logging (Fog → Cloud): Round-level validation metrics (RMSE, MAE, R^2) are recorded and plotted. Deployment (Cloud/Fog → Edge): The new global model is deployed to the edge nodes for local inference and (optional) personalization.

3.3.7 Computational and communication considerations

The design is optimized for efficiency: Edge compute: Local training is done using mini-batch optimization with a small number of epochs, making it feasible on low-end hardware. Communication: Only the model parameters are communicated. For a model with P parameters, the communication complexity per round is $O(P)$ per client, which is much lower than communicating the raw multi-year time series data. Latency: The Fog layer reduces latency compared to a cloud-only coordinator, enabling near-real-time scheduling and early diagnosis.

The proposed Edge-Fog-Cloud architecture facilitates reliable agricultural forecasting even under real-world constraints by integrating the following: Edge privacy and sovereignty (raw data never leaves the country), Fog-level federated orchestration (scalable training and aggregation), Cloud-level lifecycle services and simulation (governance, monitoring, reproducibility). The proposed architecture provides a practical foundation for deploying cross-country agricultural intelligence systems that can sustain high performance while simultaneously ensuring data confidentiality and reducing reliance on a central system.

3.4 Federated Learning Framework

This section introduces the federated learning framework developed in the proposed Edge-Fog-Cloud architecture. The framework aims to facilitate joint model training across multiple countries while maintaining data sovereignty, privacy, and resilience against non-IID data distributions.

3.4.1 Federated learning setting

In the proposed system, each participating country acts as a client in the federated learning system. Agricultural and climate data are kept locally in each country and are never shared with remote servers. Rather than sharing the data, the countries engage in joint

model training by exchanging only the model parameters. The federated learning setting is particularly relevant to this research for the following reasons: regulatory and sovereignty issues that restrict cross-country data sharing; structural heterogeneity across countries (non-IID data distributions); and differences in climate patterns, crops, and production levels. The goal is to develop a global model that identifies common temporal and climatic patterns worldwide while preserving local data sovereignty.

3.4.2 Local data preparation and sequence formation

Each country prepares supervised learning datasets from annual data, using a fixed five-year temporal window. For each prediction year, the model uses climate and agricultural data from the preceding 5 years as input to predict the yield for the next year. This sliding-window mechanism enables the model to learn about temporal relationships and the lagged effects of climate on crop yields. Local feature normalization is used to maintain numerical stability during neural network training. The normalization statistics can be computed locally or globally depending on the deployment settings. The country-level datasets vary in size based on the historical data available, leading to heterogeneous client contributions to the federated learning process. Feature normalization was performed using z-score standardization computed solely from the training partitions to prevent temporal data leakage.

3.4.3 Global model architecture

The global model used in this research is a sequence learning model. A Long Short-Term Memory (LSTM) is applied to model the temporal patterns in multi-year climate patterns. The model can learn long-term dependencies between extreme weather events and yield responses. To model multiple crop types in a single framework, crop labels can be embedded in the feature space, allowing the model to share parameters across crops while preserving crop-specific behavior. The final layer is a regression layer to predict annual yield values. The model is structured to strike a balance between expressiveness and communication efficiency, ensuring that parameter transmission between clients and the aggregator is computationally tractable.

The detailed DL-LSTM architecture is defined as follows. The model consists of an embedding layer, two stacked LSTM layers, and fully connected regression layers. Initially, the categorical crop type information was encoded as dense vector representations by using an embedding layer with an embedding dimensionality of 16. The sequential climate and agriculture data were processed by passing them through the two LSTM layers of size 128 and 64, respectively. A dropout rate of 0.2 was applied after each LSTM layer to prevent overfitting.

After that, the final LSTM layer was passed through two fully connected dense layers with sizes 64 and 32, respectively, using the ReLU activation function. Finally, the last regression output layer used a linear activation function in order to generate crop yield predictions. The model was trained using the Adam optimizer with a learning rate of 0.001 and a Huber loss function.

To achieve methodological transparency and reproducibility, the precise setup of the proposed DL-LSTM architecture and training parameters is provided in Table 4. The architecture design was chosen based on a trade-off among temporal representational

capacity, communication efficiency, and convergence stability in federated learning settings.

Hyperparameter tuning was performed through experimentation to ensure the model converged stably without being overly complex or causing excessive communication overhead when using federated optimization. In addition, the stacked LSTMs enable the model to learn long-term dependencies, and dropout helps it generalize better across non-IID distributions at the country level.

3.4.4 Federated training procedure

The federated training process involves an iterative communication process. At the start of each communication round, the Fog layer broadcasts the current global model to a chosen subset of participating countries. The selection of countries may be based on availability, connectivity, and minimum data size requirements. Each chosen country carries out local training for a few epochs using mini-batch optimization. Once local updates are complete, the countries send the updated model parameters back to the Fog layer. The Fog layer combines received updates using a weighted averaging algorithm. Countries with larger datasets get proportionally more representation in the global update to ensure statistical fairness and stability. The updated global model is then sent for the next communication round. This process continues until the convergence criteria are satisfied or a fixed number of communication rounds are completed.

To enhance convergence stability and consistency, federated learning was implemented by setting fixed communication and optimization hyperparameters across all federated learning tasks. During training, each local client optimized their model over $E = 5$ local epochs before aggregating parameters through the central node. In each communication round, a client participation ratio of $\rho = 0.8$ was set, meaning that 80% of clients at the country level participated in the aggregation process.

The federated optimization process ran for 100 communication rounds using the Adam optimizer algorithm with an initial learning rate of 0.001. A stepwise learning rate decay method was used during training, with the learning rate halved after every 25 communication rounds to facilitate convergence at the end of the training period.

To ensure robustness against outlier data and extreme climate variability, the Huber loss function was used in the federated optimization process with a $\delta = 1.0$ transition parameter. Mini-batching with a batch size of 32 was performed in all local clients' optimization processes.

Table 4 DL-LSTM hyperparameters

Hyperparameter	Value
LSTM layers	2
Hidden units	128, 64
Dropout	0.2
Dense layers	64, 32
Embedding size	16
Activation	ReLu
Output activation	Linear
Optimizer	Adam
Learning rate	0.001
Loss function	Huber
Batch size	32

The detailed hyperparameters for federated optimization used throughout the distributed training process are summarized in Table 5.

Optimal hyperparameter values were chosen empirically through careful tuning to achieve convergence stability and efficient communication, among other properties, across heterogeneous multi-country farming data distributions. For example, partial client participation and adaptive learning rates contributed to achieving federated convergence stability.

3.4.5 Handling non-IID data

One of the most important issues in multi-country federated learning is data non-IID-ness. Yield scales, climate variability, and the frequency of extreme events vary widely across countries. To counter the effects of instability caused by non-IID data, the following techniques are employed: weighted aggregation based on local dataset sizes; robust regression loss functions to make the model less sensitive to outliers; limited local epochs to avoid overdrift in local models; and optional clipping of updates to stabilize aggregation. These techniques help improve convergence stability and avoid divergence caused by highly heterogeneous local updates.

3.4.6 Communication and efficiency considerations

The federated learning system is designed to optimize communication efficiency. Only model parameters are communicated in each round, not the raw time series data. This greatly reduces the communication bandwidth compared to centralized learning. Because communication is usually the bottleneck in distributed learning, a trade-off is made between the number of local training epochs and the number of communication rounds to converge efficiently without incurring high transmission costs. Moreover, the Edge-Fog-Cloud architecture helps reduce latency by enabling aggregation at the Fog level, bypassing direct communication with the distant cloud server.

3.4.7 Centralized and edge baselines

To conduct a thorough comparison, the federated learning framework is evaluated against a fully centralized training method, where all data are combined into a single dataset, and an edge-only method, where each country trains its own separate model. This allows evaluating the trade-offs among prediction accuracy, privacy protection, and scalability. The federated learning framework presented in this work allows for joint agricultural yield prediction across multiple sovereign states without exchanging raw data. By leveraging local sequence modeling, weighted aggregation, and robustness,

Table 5 Federated learning hyperparameters

Hyperparameter	Value
Local epochs E	5
Client participation ratio p	0.8
Communication rounds	100
Optimizer	Adam
Initial learning rate	0.001
Learning rate decay	0.5 every 25 rounds
Loss function	Huber
Huber threshold δ	1.0
Batch size	32

the framework enables scalable, privacy-conscious learning in heterogeneous, non-IID environments.

3.4.8 Privacy assumptions and threat model

The proposed federated agricultural forecasting approach aims to minimize direct data sharing among participants by enabling collaborative parameter tuning. In the proposed system, data on agriculture and weather conditions will be locally maintained at edge clients, while only parameter updates will be exchanged during the federation process.

Nevertheless, the proposed system does not guarantee security against sophisticated federated learning attacks, including gradient inversion, model membership inference, model reconstruction, and other attacks that exploit parameter updates. The present work assumes an honest-but-curious federated learning setup in which participating clients and aggregation servers properly follow the training procedure but seek to derive additional insights from parameter updates.

Federated learning experimentation is primarily focused on parameter sharing via decentralized learning and on federated aggregation of updates using the FedAvg algorithm. While methods such as secure aggregation, encryption for communication channels, and differential privacy are presented theoretically as part of the architecture, they have not been experimented with in the simulated system.

Thus, claims of privacy protection within the developed architecture can only be interpreted as protection against direct data exposure. Further research could include experimenting with more rigorous privacy protection techniques, such as secure aggregation, homomorphic encryption, and differential privacy.

3.5 Model architectures

This work assesses three types of predictive models in order to provide a thorough examination of the performance of centralized, edge-only, and federated learning approaches: classical machine learning models, deep learning models, and hybrid models. The goal is to compare models with varying representational capabilities and structural assumptions under the same temporal forecasting tasks.

The choice of classical machine learning models was made to provide strong, interpretable baselines for regression tasks involving tabular, correlated climate variables. Random Forest regression was chosen for its ensemble-based architecture, tolerance to multicollinearity, and ability to capture non-linear interactions between features without requiring extensive feature engineering [39]. Moreover, gradient-boosting-based models, such as XGBoost, were used to capture complex non-linear relationships and hierarchical feature interactions, which are particularly important in climate-yield relationships where interactions among temperature, precipitation, and extreme events are not strictly linear.

For deep learning, a sequence-based LSTM architecture was used to capture the temporal dependencies over multi-year climate windows. The LSTM network is intended to capture the lagged climatic influences and the long-term temporal dependencies that are not explicitly modeled in static machine learning models [40]. The network uses a fixed five-year sliding window of input features to make a single continuous yield prediction for the following year. This setup enables the network to learn dynamic representations of climate variability and its temporal influence on crop yields.

Data sets of agricultural yields typically include both long-term climatic temporal dependencies and robust statistical feature structures in their interdependencies. Although long-short-term memory (LSTM) models in deep learning can be effective for learning temporal patterns, they might be inefficient at modeling the structure of agricultural statistical interdependencies. On the other hand, although machine learning and dense learning can be very efficient at modeling structured interdependencies, they might fail to effectively capture temporal sequence information.

To better explore the advantages of joint temporal and structured feature learning, a hybrid model architecture was developed that integrates sequential modeling with structured feature modeling. In the proposed framework, the temporal climate variable sequence is first modeled through an LSTM network to obtain temporal embeddings. The temporal sequences of climate variables, such as temperature, precipitation, and humidity, as well as past yield sequences, were modeled using the DL-LSTM network branch. At the same time, static or summary agronomic variables were modeled using fully connected layers to capture interactions among cross-sectional features. Selected agronomic and statistical indicators were forwarded to the structured dense learning network branch. The outputs from the temporal and structured modeling branches are concatenated and further processed with additional fully connected layers to produce the final regression output. The overall structure of the proposed hybrid architecture is illustrated in Fig. 10.

As illustrated in Fig. 10, the hybrid model architecture comprises two parallel branches. The temporal branch receives five-year climatic sequential inputs and employs an LSTM layer to learn dynamic temporal embeddings. Simultaneously, agronomic structural features are fed into fully connected layers to learn cross-sectional interactions. The learned embeddings from both branches are concatenated during the feature fusion phase and then fed into further dense layers to produce the final yield prediction. This allows the model to learn both temporal and structural feature relationships in a unified regression framework.

This two-branch modeling approach enables the model to learn both long-term temporal dependencies and high-order feature interactions that a sequential model alone cannot model. In agricultural systems, crop yield is determined not only by temporal climate variability but also by structural factors such as crop type, harvest region, and production scale. From a federated learning perspective, the hybrid model offers another benefit. The temporal branch helps learn generalized climate-yield patterns, and the structured branch helps learn country-specific patterns. This improves the model's robustness to non-IID data by preventing overfitting to local temporal patterns.

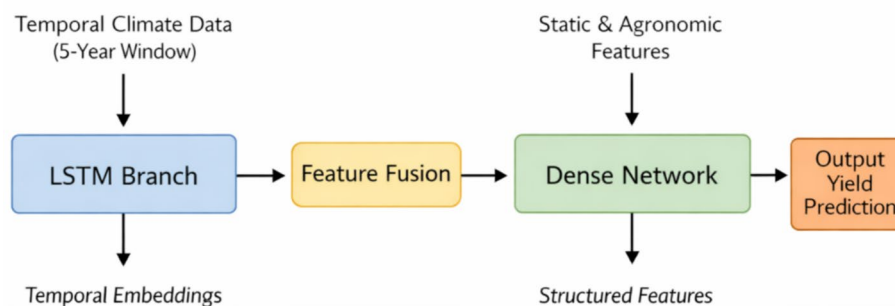


Fig. 10 Overall architecture of the proposed hybrid DL-LSTM and structured feature fusion framework

Additionally, because the model is relatively low-complexity compared to a deep transformer, communication costs are reduced during federated learning. The hybrid model serves as an intermediate between traditional machine learning and deep learning to comprehensively analyze the effect of expressiveness on federated learning.

Each model was trained and tested on the same data split and evaluation criteria. In the federated setup, the same architecture was cloned on each client, and only updates were shared during aggregation. This ensured that any differences in performance across the centralized, edge-only, and federated scenarios were due solely to the learning paradigm and not to architectural differences. The addition of diverse model families enables a thorough analysis of the impact of architectural complexity on data distribution constraints, communication, and privacy-preserving learning environments.

3.5.1 Hybrid feature fusion strategy

The proposed hybrid approach entails incorporating both temporal dependency modeling and structured statistical feature modeling within a common regression model. To promote diversity in the model and ensure robust output, the input features were divided into two distinct groups based on temporal and statistical characteristics, respectively.

Temporal input features, comprising sequential climate features, yield information and indicators of extreme climate, were modeled using the LSTM branch to capture the dynamics underlying the climate-yield association. Static agronomic features, including crop descriptors, harvested areas, production volumes, and national statistics, were used for statistical modeling in the structured, dense branch.

Let X_{seq} denote the sequential temporal input features and X_{stat} represent the structured statistical agricultural features. The temporal embedding generated by the LSTM branch is formulated as:

$$h_{DL} = LSTM(X_{seq}) \quad (8)$$

The structured feature representation produced by the dense learning branch is defined as:

$$h_{ML} = f_{Dense}(X_{stat}) \quad (9)$$

The learned latent representations extracted from both branches are fused through feature concatenation:

$$h_{fusion} = [h_{DL}; h_{ML}] \quad (10)$$

Finally, the fused representation is forwarded to the regression layer for yield estimation:

$$\hat{y} = Wh_{fusion} + b \quad (11)$$

Here, W and b represent the trainable regression coefficients. This hybridization approach allows the model to learn both long-term temporal correlations and structured agriculture relations simultaneously. To better explain the feature-splitting scheme used in the hybrid architecture, the distribution of features across learning branches is illustrated in Table 6.

This suggested approach enables the LSTM branch to focus on dependency learning, while the dense branch can exploit complementary relationships within the

Table 6 Feature partitioning strategy in the hybrid framework

Feature group	Processing branch
Historical yield sequences	LSTM branch
Temperature indices	LSTM branch
Precipitation variables	LSTM branch
Extreme climate indicators	LSTM branch
Harvest area	Dense branch
Production quantity	Dense branch
Crop descriptors	Dense branch
Country-level statistical summaries	Dense branch

Table 7 Summary of mathematical symbols and variables used in the proposed framework

Symbol	Description
$X_{t-5:t-1}$	Multi-year temporal input sequence
y_t	Annual crop yield target value at year t
D_c	Country-specific local dataset
$\theta^{(r)}$	Global model parameters at communication round r
$\theta_c^{(r)}$	Local client model parameters
$\Delta \theta_c^{(r)}$	Client model update transmitted to Fog layer
$B_{c,e}$	Mini-batch sampled from client c
η	Learning rate
E	Number of local training epochs
ρ	Client participation ratio
S_r	Selected client subset at round r
N_c	Number of samples owned by client c
X_{seq}	Sequential temporal feature inputs
X_{stat}	Structured statistical feature inputs
h_{DL}	Temporal embedding generated by LSTM branch
h_{ML}	Structured feature representation from dense branch
h_{fusion}	Concatenated hybrid feature representation
W, b	Trainable regression parameters
$\hat{y}$	Predicted crop yield value

agriculturally structured data. Such a design not only helps ensure regression stability but also explains the superior performance of the hybrid system in agricultural data environments.

The federated learning-based forecasting model comprises several mathematical equations that involve various optimization parameters for distributed learning, time-series modeling, and data fusion. To clarify and make sense of the equations presented in the paper, the major mathematical symbols and variables used in this research are tabulated in Table 7.

Table 7 depicts the major mathematical symbols and variables utilized in the proposed federated learning and hybrid forecasting framework.

3.6 Experimental setup

This section outlines the experimental setup used to test the proposed Edge-Fog-Cloud federated learning framework. The experiments are designed to compare centralized, Edge-only, and federated learning approaches on a common set of data splits and metrics. The data cover the period from 1990 to 2022 and are panel-structured. To prevent any information leakage, a time-aware split was employed. For each country, the earliest

years were used for training. The years 2015–2018 were reserved exclusively for validation and early stopping. The final years were set aside for testing, and these were used exclusively for that purpose. This temporal split ensures that future crop yield predictions are based solely on past knowledge, maintaining the realism of crop yield forecasting. In the federated learning scenario, each country has its own local train-validation split. The global model is tested on a separate aggregated test set. To test the efficacy of the federated architecture, three learning paradigms were evaluated.

The experiments were performed in a simulation environment using the Python framework, where the behaviors of decentralized client involvement, federated aggregation, and Edge-Fog-Cloud orchestration were simulated using various agricultural datasets. The experiment primarily focused on learning and optimization behaviors rather than infrastructural benchmarking.

The period from 1990 to 2014 was designated as training data, 2015 to 2018 as validation and early stopping data, and 2019 to 2022 as test data in the experiment. This temporal splitting technique was consistently used across all countries participating in the experiment.

Centralized learning: Data from all countries were combined into a single dataset. A global deep learning model was trained using mini-batch optimization. This setup is used as the baseline because the model has access to all the raw data without any privacy constraints.

Edge-only learning: Each country trained a local model on its own data. There was no parameter sharing. This setup is the most privacy-friendly but does not allow knowledge sharing across countries.

Federated learning: In the federated learning architecture, each country trained a local model on its own data. Only model parameters were shared. The Fog layer aggregated the client updates using weighted averaging. Multiple communication rounds were performed until convergence. Client participation was managed through a participation ratio to simulate a realistic distributed setting.

Although all three approaches—centralized, edge-only, and federated—were considered within the proposed methodology, not all model families have been tested under all three paradigms. Resource limitations led to using the DL-LSTM family as the backbone for all experiments in the federated and edge-only paradigms. The hybrid and classical machine learning models have been tested in centralized experiments because of their potential to account for the level of architectural expressiveness and model complexity when full data is available.

A sequence-based deep learning architecture was employed in all deep learning experiments. Input: 5-year sliding window of climate and agricultural variables. Core model: Long Short-Term Memory (LSTM). Output: Single continuous yield prediction. The LSTM layer is responsible for modeling the temporal dependencies and lagged climate impacts on agricultural yields.

The training hyperparameters employed were Optimizer: Adam, Learning rate: fixed small learning rate for stable convergence, Batch size: moderate mini-batch size, Number of local epochs (federated): restricted to avoid client drift, Number of communication rounds: multiple rounds until validation convergence, Loss function: robust regression loss (Huber loss) to reduce the effect of outliers. Early stopping was employed based on validation performance to prevent overfitting.

The performance of the models was assessed using the following regression evaluation metrics: Coefficient of Determination (R^2), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). While R^2 indicates the amount of variation explained, MAE and RMSE indicate the magnitude of errors, absolute and squared, respectively. Performance is primarily reported on the aggregated global test dataset, while country-level validation statistics were monitored during federated optimization.

In federation experiments, only suitable clients with sufficient data were included. Client selection was performed in each communication round. Aggregation took place at the Fog layer. Validation performance was monitored in each round. The number of local periods was limited to ensure stability in independent and non-identical distribution environments. Weighted aggregation was used. A robust loss function was employed. Optional update clipping was investigated. A total of 39 country-level clients satisfying the minimum historical data availability criteria participated in the federated learning experiments. Countries with insufficient temporal data to construct stable five-year sequential windows were excluded from federated optimization experiments.

All experiments were performed using Python-based deep learning libraries. TensorFlow/Keras for deep learning, Scikit-learn for machine learning, and a custom simulation of the Edge-Fog-Cloud workflow for federated orchestration. For hardware, a GPU-powered environment was used for centralized learning, and distributed simulation was used for federated learning. Repeatability was ensured by using fixed random seeds and the same preprocessing pipelines for all experimental settings. Experiments were carried out using Python code that used the TensorFlow library on the MSI GE62VR machine with an Intel Core i7 CPU, an NVIDIA GeForce GTX series graphics card, and 32 GB of RAM. All experiments were initiated by specifying the same seed value of 42 for reproducibility. Early stopping with a patience of 10 epochs was used to prevent overfitting. The Adam optimizer was set to an initial learning rate of 0.001, with the standard momentum parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$. Therefore, system-level deployment metrics such as communication latency, bandwidth utilization, aggregation overhead, and client availability fluctuations were not experimentally benchmarked in this study.

4 Results

This section describes the comparative assessment of centralized, federated, hybrid, and traditional machine learning models for multi-country cereal yield prediction. The performance of the models is measured using RMSE, MAE, and R^2 on the aggregated test dataset (see Table 8).

From the analysis of the resulting data, several interesting observations can be made. First, the edge-only DL-LSTM model performed poorly on predictive tasks, with the highest error rates, an R^2 of 0.4315, an RMSE of 1685.42, and an MAE of 1248.77. The above results clearly show that isolated country-level training is hampered by limited data variability and the inability to generalize across the heterogeneity of climate-yield distributions.

Second, the federated DL-LSTM model was significantly better at predicting, with an R^2 of 0.5222 and an RMSE of 1530.21. These results show that collaboration in the learning process, when countries share parameters, enables the model to generalize climate-yield correlations even without directly accessing agricultural data.

Table 8 Summary of predictive performance

Learning paradigm	Model	RMSE	MAE	R ²
Edge-Only	DL-LSTM	1685.42	1248.77	0.4315
Federated	DL-LSTM (FedAvg)	1530.2168	1123.5373	0.5222
Centralized	DL-LSTM	1377.4485	1057.0931	0.6128
Centralized	Hybrid (DL-Emb + ML)	1277.7145	959.7848	0.6668
Centralized	Classic ML	814.5736	590.1532	0.8643

The third finding is that the centralized DL-LSTM model outperformed the two decentralized approaches, with an R² value of 0.6128. Higher performance is anticipated since the centralized approach directly accesses the whole set of multi-country data during training.

Further improvement in performance was observed with the hybrid model, which yielded an R² of 0.6668 and an RMSE of 1277.71. These results confirm that combining time-series modeling and structured feature learning increases regression stability and effectiveness.

Lastly, the highest performance levels were observed with the centralized Classical ML model, with an R² of 0.8643, RMSE of 814.57, and MAE of 590.15. These results confirm the high applicability of ensembles-based machine learning algorithms for complex, structured agricultural datasets.

Figure 11 shows the true vs. predicted scatter plots for edge-only, federated, centralized deep learning, hybrid, and classical machine learning models.

The classical ML model is closely clustered around the diagonal line, signifying a good predictive fit. The hybrid model depicts better control over dispersion than the LSTM model. The centralized LSTM model has high variance, especially in high-yielding regimes.

The DL-LSTM algorithm that uses only the edge model shows the largest prediction deviation and the least resemblance to the desired prediction curve because of the isolated optimization process within each country and the lack of exposure to diverse climate-yield relationships. On the other hand, the federated DL-LSTM model shows significantly better clustering performance and lower prediction deviation due to parameter sharing across distributed country nodes.

To better understand the training dynamics of the federated learning framework, the validation performance over the communication rounds is investigated. Figure 12 shows the validation RMSE and R² values over the federated optimization process.

As evident from Fig. 12, the federated model converges steadily round by round. The validation RMSE reduces steadily in the initial rounds, signifying efficient knowledge distillation from the decentralized country-level clients. There is a swift convergence in performance in the initial 40–60 rounds, followed by a gradual stabilization phase. Unlike federated learning, the edge-only model allows for fully isolated optimizations that are not reliant on any parameter-sharing or communication iterations. Consequently, it is impossible to observe convergence in an edge-only environment.

To further explore the effect of heterogeneous, non-IID data distributions at the country level, an additional analysis of the countries' performance was conducted. The objective of this analysis is to examine the effects of varying historical sample sizes, climatic variation, and local agriculture on the performance of the federated learning process.

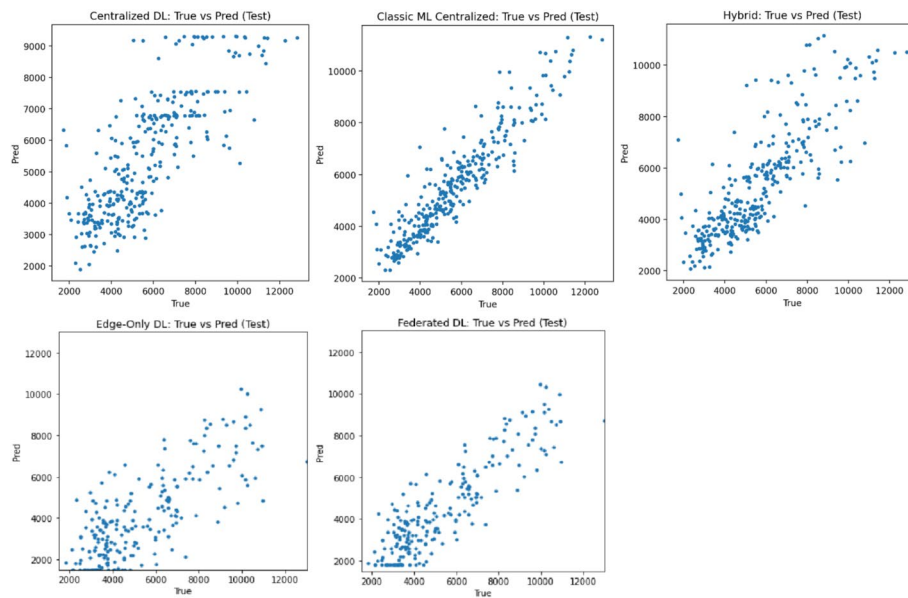


Fig. 11 Comparison of true and predicted crop yield distributions across edge-only, federated, centralized deep learning, hybrid, and classical machine learning models

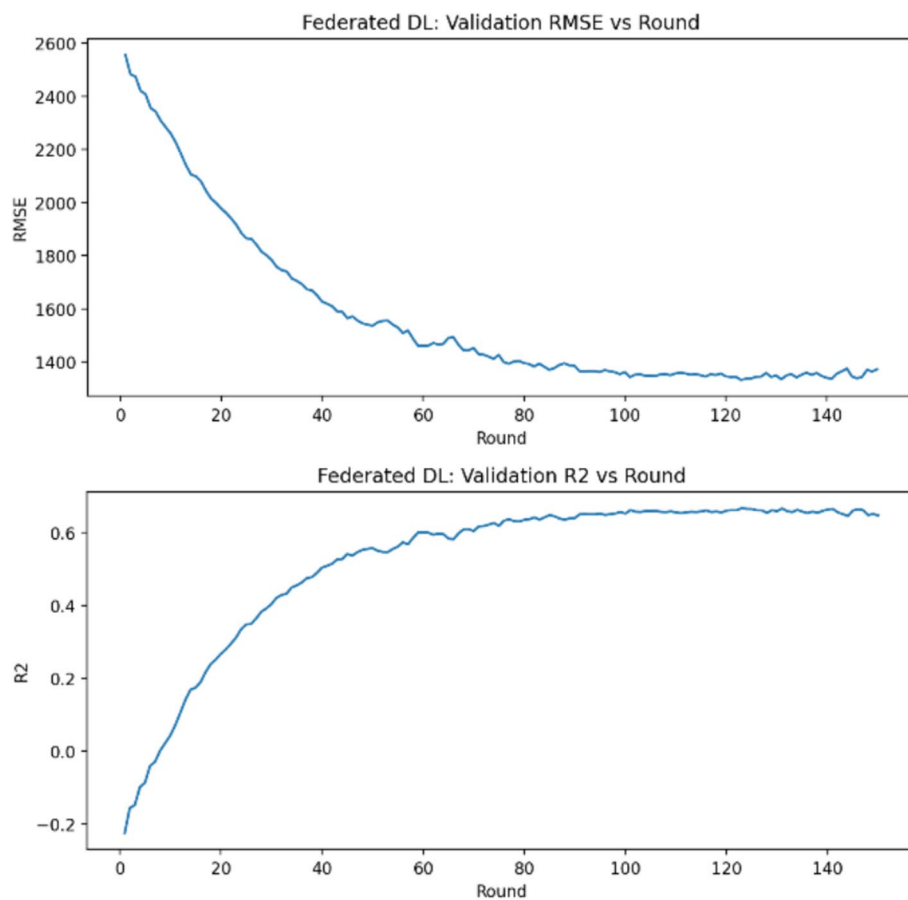


Fig. 12 Evolution of validation RMSE and R^2 across federated communication rounds. Since the edge-only paradigm does not involve collaborative optimization or communication rounds, convergence analysis is only reported for the federated learning framework

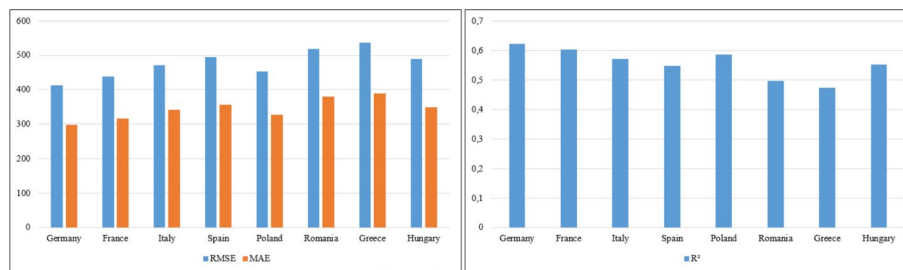


Fig. 13 Representative country-wise federated DL-LSTM forecasting performance under heterogeneous non-IID agricultural data distributions

Figure 13 shows the country-level performance for the federated DL-LSTM model in selected European agricultural clients.

As shown in Fig. 13, the forecasting results from the federated forecasting approach varied across several representative participants, depending on the amount of historical data available, climate variability, and the spatial heterogeneity of agriculture within each country. For example, countries with more observations showed better predictive consistency than those with fewer observations or high climate variability.

Concurrently, the validation R^2 value steadily improves from around 0.65, indicating that the model learns relevant global representations despite the non-IID country-level data distribution. After the 100th round of communication, the incremental improvements in performance become negligible, indicating convergence of the global model. These results confirm that federated averaging successfully distills decentralized updates into a unified global forecasting model. These results suggest that the proposed federated optimization framework maintains stable convergence behavior despite heterogeneous country-level data distributions.

5 Discussion

The results of this research offer several key takeaways regarding distributed agricultural forecasting tasks under privacy and heterogeneity constraints. First, the effectiveness of centralized classical machine learning models attests to the continued robustness of ensemble-based tree models for structured and tabular agricultural data. The R^2 value of 0.8643 indicates an efficient representation of non-linear relationships among climate variables, agronomic factors, and agricultural yield, without the need for deep sequence modeling. This aligns with previous studies showing that gradient-boosting-based models are generally superior to deep learning models for medium-sized tabular data.

Second, the performance difference between centralized deep learning models and classical machine learning models highlights an important characteristic of the data. The annual granularity and short five-year horizon make it difficult to leverage the benefits of long-range sequence modeling. Although the LSTM model is successful at modeling temporal relationships, the benefit of doing so is not as high as that of tree-based feature interaction models. This indicates that, for annual agricultural forecasting tasks, feature interactions may play a more important role than long-term temporal memory. These findings imply that the agricultural data used in this study exhibit strong statistical dependencies that are better modeled by ensemble techniques than by sequential deep learning methods alone.

Third, the hybrid approach demonstrates that combining learned temporal embeddings with structured regression is more effective than deep learning alone. The combination of LSTM embeddings with non-linear regression improves stability and decreases dispersion, especially for mid-to-high yield regimes. Nevertheless, the hybrid approach fails to outperform the classical ML baseline, suggesting that the representational advantages of deep embeddings do not entirely offset the statistical efficiency advantages of ensemble methods in structured data problems.

Most significantly, the federated learning architecture demonstrates that moderate predictive performance can be achieved compared with centralized deep learning. As observed, classical ML algorithms perform better than federated deep learning approaches on an annual agricultural panel dataset, suggesting that the former is superior to the latter for this type of dataset. However, the main aim of the federated DL-LSTM approach proposed herein is not only to attain high prediction accuracy but also to explore privacy-preserving collaborative forecasting in decentralized agricultural environments. Despite poor performance relative to the centralized deep learning model ($\Delta R^2 \approx 0.09$), the decrease in predictive power remained limited despite considerable country-level data heterogeneity and stringent requirements for decentralized data.

Moreover, validation behavior at the country level varied across federated clients due to differences in factors such as climatic variability, production scale, and historical data availability. Countries with more consistent past agricultural output had lower local prediction errors, while those with highly variable production levels had higher local prediction errors in federated optimization.

The edge-only training approach highlighted the shortcomings of independent local optimization in highly diverse agricultural settings. Due to independent training in each nation with no parameter transfer, these models lacked generalization and adequate training across broader climate-crop dynamics. The performance difference between edge-only and federated learning systems demonstrates the advantages of collaborative distributed optimization, even with raw agricultural data localized within nations.

The convergence curves also reveal the stable optimization of federated learning. The rapid progress followed by slow convergence indicates that the global model is learning the climate-yield patterns common to all countries before learning the remaining country-specific variations. The point of diminishing returns occurring at 100 communication rounds also reveals opportunities to optimize training cost through communication-efficient early-stopping methods.

From a systems perspective, the experiments above demonstrate the feasibility of the proposed Edge-Fog-Cloud model. The Edge-Fog-Cloud model successfully balances three conflicting demands of predictive accuracy, privacy preservation, and communication efficiency. While the classical machine learning approach offers the best predictive accuracy, the federated learning approach can address multi-country collaboration challenges where direct data sharing is not feasible.

These results suggest that the decision between centralized and federated learning for agricultural forecasting should not be based solely on predictive performance. Rather, the system design should account for governance, scalability, and communication constraints. In practical cross-border agricultural intelligence systems, federated learning provides a practical compromise between performance and sovereignty.

6 Conclusion

This work introduced a privacy-preserving, scalable agricultural yield forecasting framework based on an integrated Edge-Fog-Cloud architecture. The proposed framework enables collaborative multi-country learning with strict data sovereignty, leveraging federated learning techniques. We used a multi-decade European cereal dataset to systematically compare centralized, edge-only, and federated learning approaches with consistent temporal evaluation conditions.

The experimental results show that classical ensemble machine learning approaches achieve the best predictive performance in a centralized setup, underscoring the effectiveness of classical machine learning for modeling structured tabular agricultural data. Deep learning approaches with LSTM architectures offer useful temporal modeling power but do not outperform classical approaches in annual-resolution panel data. The hybrid model, combining temporal embeddings with structured regression, improves on the deep learning model, highlighting the advantage of joint sequential and structural modeling.

More importantly, the empirical results indicate a distinct performance ranking among edge-only, federated, and centralized learning systems. Whereas the edge-only system faced generalization problems due to local-only training, the federated learning model achieved significant improvements in prediction accuracy through collaborative parameter updates without sharing any agricultural data. Despite centralized learning providing the best deep learning performance, the federated model proved effective in predicting trends while safeguarding data privacy.

Despite the promising results, the proposed architecture still has some weaknesses. Firstly, federated deep learning performed somewhat poorly in terms of prediction due to the strong non-IID properties of multinational agricultural datasets. Secondly, the yearly temporal frequency and the relatively short sequence length may hinder LSTMs' ability to capture climate yield dependencies at a larger scale. Thirdly, federated learning adds another layer of overhead due to communication costs and the increased complexity of convergence associated with federated optimization. Additionally, this research was conducted using a European cereal production data set, potentially limiting the generalizability of the framework when applied to other agricultural areas or climate data sets with different frequencies. Lastly, personalization methods and adaptive aggregations are not yet integrated into this framework. In addition, the proposed framework was evaluated in a simulation-based federated environment, without real deployment-level benchmarking, to assess communication latency, aggregation overhead, bandwidth consumption, and dynamic client availability. Moreover, the existing framework lacks experimental validation of adversarial privacy attacks, including gradient inversion, membership inference, and parameter update-based model reconstruction. Furthermore, the present study fails to experimentally explore the federated implementation of ensemble machine learning algorithms, such as federated Random Forests and federated gradient boosting.

Future research can address the shortcomings mentioned above by leveraging personalized federated learning approaches, adaptive client selection methods, efficient communication aggregation processes, and transformer-based temporal models capable of capturing long-range climate dependencies. Additional improvements can be achieved by incorporating high-frequency climate and remote sensing data, and by adopting

hierarchical regional federated learning aggregation approaches. Future studies may further extend the proposed framework toward real deployment-oriented Edge–Fog–Cloud infrastructures involving communication-aware federated optimization, adaptive client scheduling, latency analysis, and large-scale distributed orchestration benchmarking. Further research might extend the suggested model by incorporating federated tree-based ensemble learning, federated gradient boosting, and federated personalized learning methods in a heterogeneous agricultural setting.

Author contributions

A.U.: Conceptualization, methodology, data curation, writing–original draft. A.B.: Literature review, investigation, validation. H.A.: Data collection, resources, validation. M.D.H.: Supervision, methodology, writing–review and editing. A.A.: Project administration, supervision, writing–review and editing. All authors have read and agreed to the published version of the manuscript.

Funding

The authors extend their appreciation to Princess Nourah bint Abdulrahman University for funding this research through the Researchers Supporting Project (PNURSP2026R548), Riyadh, Saudi Arabia.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Declaration of the use of AI

The authors used AI-assisted language tools only for limited grammar refinement and language improvement. All scientific interpretations, methodological developments, analyses, and final manuscript revisions were performed and verified by the authors.

Competing interests

The authors declare no competing interests.

Author details

¹Computer Engineering Department, Faculty of Engineering, Munzur University, Tunceli 62000, Turkey

²Faculty of Education and Liberal Arts, INTI International University, 71800 Nilai, Negeri Sembilan, Malaysia

³Department of Accounting and Taxation, Siverek Vocational School, Harran University, Şanlıurfa, Turkey

⁴ARUCAD Research Centre, Arkin University of Creative Arts and Design (ARUCAD), Kyrenia, Northern Cyprus, Mersin 10, Türkiye

⁵Department of Economics, College of Business Administration, Princess Nourah Bint Abdulrahman University, P. O. Box 84428, 11671 Riyadh, Saudi Arabia

⁶Department of Economics, College of Business, Islamic University of Madinah, Madinah, Saudi Arabia

Received: 8 March 2026 / Accepted: 16 June 2026

Published online: 21 June 2026

References

1. Mutale B, Dai S, Chen Z, Maulu S. Enhancing food security amid climate change: assessing impacts and developing adaptive strategies. *Cogent Food Agric.* 2025;11(1):2519800.
2. Celis, J., Xiao, X., Wagle, P., Adler, P.R., White, P. (2024). A Review of Yield Forecasting Techniques and Their Impact on Sustainable Agriculture. In: Melkonyan-Gottschalk, A., Daus, D. (eds) *Transformation Towards Circular Food Systems. Greening of Industry Networks Studies*, vol. 13. Springer, Cham https://doi.org/10.1007/978-3-031-63793-3_8
3. Javed MA, Murad MAA. Crop yield prediction in agriculture: a comprehensive review of machine learning and deep learning approaches, with insights for future research and sustainability. *Heliyon.* 2024. <https://doi.org/10.1016/j.heliyon.2024.e40836>.
4. Tyagi NK, Tyagi K, Sharma G. Self-sovereign identity-based inter-country goods movement: a conceptual framework. *SN Computer Science.* 2026;7(1):127.
5. Cravero A, Bustamante A, Negrier M, Galeas P. Agricultural big data architectures in the context of climate change: a systematic literature review. *Sustainability.* 2022;14(13):7855.
6. Lu Z, Pan H, Dai Y, Si X, Zhang Y. Federated learning with non-iid data: a survey. *IEEE Internet Things J.* 2024;11(11):19188–209.
7. Trigka M, Dritsas E. Edge and cloud computing in smart cities. *Fut Internet.* 2025;17(3):118.
8. Li Q, Wen Z, Wu Z, Hu S, Wang N, Li Y, et al. A survey on federated learning systems: vision, hype and reality for data privacy and protection. *IEEE Trans Knowl Data Eng.* 2021;35(4):3347–66.

9. Savazzi S, Nicoli M, Bennis M, Kianoush S, Barbieri L. Opportunities of federated learning in connected, cooperative, and automated industrial systems. *IEEE Commun Mag.* 2021;59(2):16–21.
10. Yu P, Teng F, Zhu W, Shen C, Chen Z, Song J. Cloud-edge-device collaborative computing in smart agriculture: architectures, applications, and future perspectives. *Front Plant Sci.* 2025;16:1668545.
11. Gong R, Zhang H, Li G, He J. Edge computing-enabled smart agriculture: technical architectures, practical evolution, and bottleneck breakthroughs. *Sensors.* 2025;25(17):5302.
12. Dembani R, Karvelas I, Akbar NA, Rizou S, Tegolo D, Fountas S. Agricultural data privacy and federated learning: A review of challenges and opportunities. *Comput Electron Agric.* 2025;232:110048.
13. Mohammed B. A comprehensive overview of federated learning for NextGeneration smart agriculture: current trends, challenges, and future directions. *Informatica.* 2025;49(1):6764.
14. Killeen P, Lin C, Li F, Kiringa I, Yeap T. IoT-based smart farming architecture using federated learning: a nitrous oxide emission prediction use case. *ACM J Comput Sustain Soc.* 2025;3(2):1–38.
15. Hiremani V, Devadas RM, Sapna R, Sowmya T, Gujjar P, Rani NS, et al. Federated learning for crop yield prediction: a comprehensive review of techniques and applications. *MethodsX.* 2025;14:103408.
16. Dey T, Bera S, Mukherjee A, De D, Buyya R. Flyer: federated learning-based crop yield prediction for agriculture 5.0. *IEEE Transact Artif Intellig.* 2025;6(7):1943–52.
17. Mukherjee A, Buyya R. Federated learning architectures: a performance evaluation with crop yield prediction application. *Softw Pract Exp.* 2025;55(7):1165–84.
18. Nirosha U, Vennila G. Enhancing crop yield prediction for agriculture productivity using federated learning integrating with graph and recurrent neural networks model. *Expert Syst Appl.* 2025;289:128312.
19. Cowlishaw R, Longépé N, Riccardi A. Balancing centralisation and decentralisation in federated learning for Earth Observation-based agricultural predictions. *Sci Rep.* 2025;15(1):10454.
20. Saxena D, Gupta I, Gupta R, Singh AK, Wen X. An AI-driven VM threat prediction model for multi-risks analysis-based cloud cybersecurity. *IEEE Transact Syst Man Cybern Syst.* 2023;53(11):6815–27.
21. Saxena D, Gupta R, Singh AK, Vasilakos AV. Emerging vm threat prediction and dynamic workload estimation for secure resource management in industrial clouds. *IEEE Trans Autom Sci Eng.* 2023;21(4):5837–51.
22. Gupta R, Saxena D, Gupta I, Makkar A, Singh AK. Quantum machine learning driven malicious user prediction for cloud network communications. *IEEE Netw Lett.* 2022;4(4):174–8.
23. Gupta K, Saxena D, Gupta R, Singh AK. Maids: malicious agent identification-based data security model for cloud environments. *Clust Comput.* 2024;27(5):6167–84.
24. Gupta K, Saxena D, Gupta R, Kumar J, Singh AK. Fedmup: Federated learning driven malicious user prediction model for secure data distribution in cloud environments. *Appl Soft Comput.* 2024;157:111519.
25. Hari P, Singh MP. Adaptive knowledge transfer using federated deep learning for plant disease detection. *Comput Electron Agric.* 2025;229:109720.
26. Piccialli F, Della Bruna C, Chiaro D, Qi P, Savoia M. AGRIFOLD: AGRICulture federated learning for optimized leaf disease detection. *Expert Syst Appl.* 2025;289:128371.
27. Cristea AM, Dobre C. Federated transfer learning for tomato leaf disease detection using neuro-graph hybrid model. *AgriEngineering.* 2025;7(12):432.
28. Choubey S, Divya D. Lightweight federated transfer learning for plant leaf disease detection and classification across multiclient cross-silo datasets. *BIO Web Conf.* 2024;82:05018.
29. Muntaqim MZ, Smrity TA. Federated learning framework for brain tumor detection using MRI images in non-IID data distributions. *J Imag Inform Med.* 2025;38(6):3909–27.
30. Muntaqim MZ, Kafi HM, Smrity TA. Privacy-aware plant disease detection: federated learning with homomorphic encryption on image data. *Vis Comput.* 2026;42(1):18.
31. Muntaqim MZ, Smrity TA, Kafi HM, Miah ASM, Al Farid F, Karim HA, et al. Federated learning meets few-shot learning: a voting ensemble based combined approach to cauliflower leaf disease classification across non-iid data distributions. *Array.* 2025. <https://doi.org/10.1016/j.array.2025.100516>.
32. <https://www.kaggle.com/datasets/kostask200/eu-cereal-yield-and-climate-data-19902022/data>.
33. McMahan B, Moore E, Ramage D, Hampson S, y Arcas BA. Communication-efficient learning of deep networks from decentralized data. *Artif Intell Stat.* 2017;54:1273–82.
34. Beaulieu JM, O'Meara BC. Navigating "tip fog": embracing uncertainty in tip measurements. *Evolution.* 2025;79(7):1131–42.
35. Aga DT, Chintanippu R, Mowri RA, Siddula M. Exploring secure and private data aggregation techniques for the internet of things: a comprehensive review. *Discov Internet Things.* 2024;4(1):28.
36. Kansara M. Cloud migration strategies and challenges in highly regulated and data-intensive industries: a technical perspective. *Int J Appl Mach Learn Comput Intellig.* 2021;11(12):78–121.
37. Venčkauskas A, Toldinas J, Morkevičius N, Serkovas E, Krištaponis M. Enhancing the resilience of a federated learning global model using client model benchmark validation. *Electronics.* 2025;14(6):1215.
38. Srichandan SK, Majhi SK, Jena S, Mishra K, Bhat R. A secure and distributed placement for quality of service-aware IoT requests in Fog-Cloud of Things: a novel joint algorithmic approach. *IEEE Access.* 2024;12:56730–48.
39. Al-Jamimi HA. Synergistic feature engineering and ensemble learning for early chronic disease prediction. *IEEE Access.* 2024;12:62215–33.
40. Vo TQ, Kim SH, Nguyen DH, Bae DH. LSTM-CM: a hybrid approach for natural drought prediction based on deep learning and climate models. *Stoch Environ Res Risk Assess.* 2023;37(6):2035–51.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.