UNIVERSITY OF DHAKA

STPS-NET: A Federated Spatio-Temporal Plant Specific Network for Yield Prediction and Recommendation in Urban Agriculture

by

Exam Roll: 2324025

Registration No: 2019-117-810, Session: 2019-10

Exam Roll: 2324002

Registration No: 2019-917-830, Session: 2019-20

Bachelor of Science

Declaration Form

University of Dhaka

We, Ahad Bin Islam Shoeb and Afser Adil Olin, thus assert that the findings of our comprehensive inquiry, conducted under the direction of Dr. Md. Mamun-Or-Rashid, Professor in the Department of Computer Science and Engineering at the University of Dhaka, are encapsulated in this project. We further assert that no segment of this project has been or is currently being submitted for the conferment of any degree or certificate anywhere.

Ahad Bin Islam Shoeb
Candidate 1

Afser Adil Olin
Candidate 2

Dr. Md. Mamun-Or-Rashid
Professor
Department of Computer Science and Engineering
University of Dhaka
Supervisor

UNIVERSITY OF DHAKA

Abstract

Smart urban agriculture is becoming increasingly important in densely populated places where land is limited and sustainable production is essential. Our work introduces STPS-NET (Spatio-Temporal Plant Specific-Network), a novel architecture for accurate yield prediction and dynamic recommendations using plant-specific neural networks. STPS-NET combines plant-type embeddings with multi-scale sensor data, such as temperature, humidity, pH, and key nutrients (N, P, K), to understand the intricate interdependencies that drive plant growth. The design extracts spatial characteristics using parallel convolutional layers with varying kernel sizes, while temporal dynamics are represented by a bidirectional LSTM block. An attention method is used to highlight crucial growth stages, ensuring that the most important developmental stages are adequately emphasised. STPS-NET's dual-branch design predicts crop yield and identifies ideal parameter thresholds for actionable recommendations, allowing for precise resource management. Furthermore, the use of federated learning enables decentralised model training across numerous urban gardens, maintaining data privacy while adjusting to a variety of environmental conditions and plant types. STPS-NET outperforms baseline models on rooftop-garden datasets, resulting in decreased mean absolute error and tailored suggestions for better resource use. This work presents a strong, privacy-preserving architecture that advances decision-support systems for modern urban agriculture.

Contents

List of Figures

List of Tables

List of Algorithms

Abbreviations

FL	**F**ederated **L**earning
STPS-NET	**S**patio **T**emporal **P**lant **S**pecific **Net**work
ML	**M**achine **L**earning
NPK	**N**itrogen, **P**hosphorus, and **K**alium (Potassium)
pH	**p**otential of **H**ydrogen
ReLU	**R**ectified **L**inear **U**nit
AI	**A**rtificial **I**ntelligence
CNN	**C**onvolutional **N**eural **N**etwork
IoT	**I**nternet of **T**hings
GIS	**G**eographic **I**nformation **S**ystem
GPS	**G**lobal **P**ositioning **S**ystem
Y	**Y**ield

Chapter 1

Introduction

Urban farming—especially rooftop gardening—has become a sustainable answer for local food production as urbanisation rises and landmass declines. Rooft gardens efficiently employ vacant areas to grow food in highly populated cities like Dhaka, where land is limited, therefore benefiting the local economy and the surroundings. Although large-scale farms have always used precision farming methods, their ideas can be modified to fit urban agriculture so improving sustainability and efficiency.

We provide in this work **STPS-NET**, a new plant-specific neural network architecture for dynamic recommendation and yield prediction. With plant-specific embeddings, STPS-NET combines sensor data—collected via soil moisture, pH, temperature, humidity, nutrient (N-P-K), and light sensors attached to ESP-32 microcontrollers—with accurate prediction of crop output and provide actionable insights for enhancing growing conditions. Built on Spatio-Temporal prediction systems that examine geographical and temporal patterns in the data, the model is fundamentally

Our method's main novelty is adding **Federated Learning (FL)** with STPS-Network, which lets individual rooftop gardens teach local models on their own private data. Periodically pooled on a central server, these local models update a strong global model by using collective insights from various urban settings, while preserving data privacy.

Along with consistent suggestions for maximising plant development, experimental results from our federated learning implementations show notable increases in yield forecast accuracy. Providing a viable route for sustainable, data-driven food production in cities, STPS-NET generally links modern precision agricultural technologies with urban agriculture.

1.1 Background and Motivation

Urbanisation is quickly changing the world and creating densely populated places where limited land makes conventional farming progressively impossible. In response, rooftop gardening has become a potential option that not only supports the local production of fresh produce—including herbs, vegetables, and ornamental plants—but also improves environmental sustainability. Rooftop gardens boost city aesthetics, help reduce urban heat islands, improve air quality, and enhance biodiversity. Furthermore, these green areas assist in controlling stormwater runoff and lessen building cooling needs, thereby lowering energy consumption and reducing the risk of urban flooding.

In this study, we present **STPS-NET** (Spatio-Temporal Plant Specific-Network), a novel design employing plant-specific neural networks for dynamic recommendation generation and precise yield prediction in urban rooftop gardens. STPS-NET captures the complex interactions that govern plant growth by combining multi-scale sensor data—such as temperature, humidity, pH, and essential nutrient levels (nitrogen, phosphorus, potassium)—with plant-type embeddings. While a bidirectional Long Short-Term Memory (LSTM) block models the temporal dynamics, the architecture uses parallel convolutional layers with varying kernel sizes to extract spatial features. An attention mechanism is incorporated to highlight important growth phases, ensuring that critical developmental stages receive focused analysis.

STPS-NET is unique in its dual-branch architecture. One branch learns ideal parameter thresholds for every plant type, while the other forecasts crop yield. This dual capability enables the system to offer customized, practical advice for resource management. These recommendations are presented through an interactive UI dashboard to enhance usability, allowing urban farmers to visualize and apply the model's insights effectively.

Moreover, STPS-NET is integrated with a federated learning architecture, enabling distributed training over multiple rooftop gardens. This approach not only safeguards data privacy by keeping raw data local but also strengthens the model by allowing it to learn from the varied environmental conditions present in different metropolitan settings. Consequently, the global model is more adept at accommodating diverse plant types, including decorative vegetation, herbs, and vegetables.

1.2 Research Objectives

The primary objectives of this research are as follows:

1. **Creating Datasets:**
    - Utilize sophisticated sensors measuring soil moisture, pH, temperature, humidity, N-P-K levels, and light intensity to gather comprehensive datasets from urban rooftop gardens.
    - Collect data from a variety of plants—including vegetables, eggplants, and decorative plants—to understand the optimal conditions required for each species.
2. **Model Creation:**
    - Develop and implement a neural network architecture that integrates bidirectional LSTM, attention mechanisms, and multi-scale spatial feature extraction via convolutional layers.
    - Incorporate plant-type embeddings to enable the model to learn environmental responses and species-specific growth patterns.
3. **Integration of Federated Learning:**
    - Establish a federated learning system to enable distributed training across multiple rooftop gardens, thereby preserving data privacy and enhancing model generalizability in varied urban environments.
    - Ensure secure communication and efficient aggregation of local model updates without sharing raw data.
4. **Advisory Suggestion System and UI Dashboard:**

- Design a dynamic recommendation system that leverages the dual-branch architecture of STPS-NET to provide actionable insights for optimal resource management.
- Develop an interactive UI dashboard to visually present these recommendations, facilitating informed decision-making by urban farmers.

5. **Performance Assessment:**

- Conduct rigorous evaluation of the proposed framework using both qualitative and quantitative metrics (e.g., mean absolute error in yield estimates, resource optimization indices).
- Compare the performance of STPS-NET with baseline models in real-world rooftop garden environments.

1.3 Research Challenges

Implementing STPS-NET in a real-world urban agriculture setting presents several challenges:

- **Data Heterogeneity and Quality:** Urban rooftop gardens exhibit considerable variability in environmental conditions, sensor types, and data quality. Addressing these discrepancies, ensuring proper sensor calibration, and managing missing or noisy data are critical for reliable model performance.
- **Model Complexity and Scalability:** Capturing the non-linear and complex relationships between multiple environmental variables and plant growth necessitates sophisticated neural network architectures. The model must be computationally efficient to run on limited hardware and scalable to handle data from an increasing number of rooftop gardens.
- **Federated Learning Constraints:** Decentralized training requires robust systems for secure data aggregation and efficient model convergence. Key technical challenges include managing communication overhead, ensuring data privacy, and addressing potential data imbalances across different nodes.

- **Integration with Existing Urban Systems:** The proposed system must seamlessly integrate with current urban farming practices and technologies, including the delivery of user-friendly outputs via a UI dashboard and compatibility with diverse sensor equipment.

- **Evaluation in Dynamic Environments:** Urban agricultural environments are inherently variable, making it challenging to develop a model that consistently performs well under fluctuating conditions. Continuous evaluation and iterative refinement based on real-world feedback are essential.

1.4 Paper Organisation

This paper is organized as follows:

- **Chapter 2** reviews the pertinent literature, providing the theoretical basis and highlighting recent developments in urban agriculture and precision farming.

- **Chapter 3** details the design and implementation of the STPS-NET architecture, including the development of the recommendation system and the integration of federated learning.

- **Chapter 4** describes the experimental setup, data collection procedures, and evaluation metrics, followed by an in-depth analysis of results obtained from rooftop garden datasets.

- **Chapter 5** concludes the paper with a discussion of the findings, practical implications, and directions for future research.

This work aims to provide a robust, scalable, and privacy-preserving framework that advances urban rooftop gardening, thereby enhancing local production, optimizing resource management, and promoting urban sustainability through modern decision-support systems.

Chapter 2

Literature Review

Recent advancements in precision agriculture have leveraged sensor technologies and machine learning methods to optimize crop management practices. This literature review synthesizes recent studies from 2022 to 2024 that utilize sensor data and machine learning, including Federated Learning, to improve agricultural outcomes.

2.1 Precision Agriculture Monitoring and IoT Integration

Chen et al.[5] developed a wireless sensor network for real-time monitoring of agricultural fields. This system employed machine learning models to predict soil moisture and nutrient levels, thus optimizing irrigation and fertilization practices. This study highlights the effectiveness of integrating IoT with machine learning to enhance agricultural productivity by providing timely and accurate data-driven insights.

Wang et al. [43] presented an IoT-based smart irrigation system that utilized machine learning algorithms to predict soil moisture levels and adjust water supply accordingly. The system ensured efficient water use and optimal crop growth, demonstrating the practical benefits of smart irrigation in precision agriculture.

Huang et al.[11] discussed the integration of IoT technologies in smart farming, focusing on the challenges and future research directions. The study emphasized the role of machine learning in analyzing the vast amounts of data generated by IoT devices, which is crucial for making informed decisions in agricultural management.

2.2 Machine Learning for Crop Monitoring and Disease Detection

Zhu et al.[47] provided a comprehensive review of deep learning techniques applied to precision agriculture, emphasizing the use of convolutional neural networks (CNNs) for image-based plant disease detection and crop yield prediction. This review underscores the growing importance of deep learning in enhancing the accuracy and efficiency of agricultural monitoring systems.

Kumar et al.[23] explored the use of unmanned aerial vehicles (UAVs) equipped with multispectral cameras to collect crop images. Machine learning models were used to analyze these images for early detection of crop diseases, showcasing the potential of UAVs and machine learning in proactive crop health management.

Lee et al.[25] developed a deep learning model for diagnosing crop diseases using image data collected from agricultural fields. The model achieved high accuracy in identifying various plant diseases, further validating the efficacy of deep learning in precision agriculture.

2.3 Federated Learning for Data Privacy and Model Improvement

Li et al.[26] implemented a Federated Learning framework to aggregate crop yield prediction models trained on data from different geographic locations. This approach preserved data privacy while improving model accuracy through collaborative learning. The study highlights the potential of Federated Learning to enhance agricultural data analysis without compromising privacy.

Ahmed et al.[2] integrated edge computing with Federated Learning to process agricultural data in real-time. The system reduced latency and improved the responsiveness of smart farming applications, demonstrating the practical benefits of combining edge computing with Federated Learning in precision agriculture.

Smith et al.[38] explored the use of Federated Learning to develop climate-resilient agricultural models. The models were trained on data from diverse climatic regions to enhance their robustness and adaptability, emphasizing the role of Federated Learning in addressing climate variability in agriculture.

2.4 AI-Driven Soil and Crop Management

Patel et al.[30] utilized a network of soil sensors and AI algorithms to monitor soil health parameters such as pH, moisture, and nutrient levels. The system provided real-time recommendations for soil management to improve crop yield, highlighting the practical applications of AI in soil health monitoring.

Khan et al.[21] presented a smart fertilization system that used IoT sensors to monitor soil nutrient levels and AI algorithms to optimize fertilization schedules. This approach improved crop yield and reduced environmental impact, demonstrating the environmental and economic benefits of AI-driven fertilization systems.

Garcia et al.[7] developed a sensor fusion system that combined data from multiple types of sensors to provide comprehensive insights into crop health and environmental conditions. AI models were used to analyze the data and support decision-making, underscoring the importance of sensor fusion and AI in precision agriculture.

2.5 Predictive Analysis and Decision Support Systems

Singh et al.[36] reviewed various machine learning techniques used for predictive analysis in agriculture, including crop yield prediction, weather forecasting, and pest detection. The study provided a detailed overview of the current state of machine learning in agricultural prediction, highlighting its potential to transform decision-making in agriculture.

Nguyen et al.[29] proposed a hybrid machine learning approach that combined traditional regression models with neural networks to improve crop yield prediction accuracy. This hybrid approach demonstrated the advantages of integrating multiple machine learning techniques to enhance predictive performance in agriculture.

Jones et al.[17] reviewed current trends in AI-powered precision agriculture, including the use of machine learning and Federated Learning to enhance crop management practices. The authors discussed the future prospects of AI in agriculture, emphasizing the need for continued innovation and research in this field.[20]

The recent studies reviewed highlight the significant advancements in precision agriculture through the integration of sensor data, machine learning, and Federated Learning. These technologies have demonstrated their potential to enhance crop monitoring, disease detection, soil management, and predictive analysis, thereby transforming agricultural practices. The integration of IoT and AI in precision agriculture not only improves productivity and efficiency but also addresses challenges related to data privacy, climate variability, and environmental sustainability. Continued research and innovation in these areas are essential to fully realize the benefits of smart farming technologies.

2.6 Use of Sensor Technologies in Precision Agriculture

Several studies have highlighted the effectiveness of sensor technologies in precision agriculture. For example, Smith et al. [37] developed an IoT-based system for real-time monitoring of soil moisture and nutrient levels, demonstrating significant improvements in crop yield. Similarly, Johnson et al. [13] utilized soil pH sensors and found that maintaining optimal pH levels can enhance nutrient uptake and crop growth.

2.7 Application of ML and FL in Agricultural Practices

Machine learning and federated learning have been increasingly applied to agricultural practices to enhance decision-making and predictive capabilities. For instance, Brown et al. [4] used ML models to predict crop yield based on environmental parameters, achieving high accuracy in diverse conditions. Additionally, Jones et al. [17] demonstrated that FL could enhance data privacy and model robustness by aggregating locally trained models from multiple farms.

2.8 Mathematical Relationships Between Environmental Parameters and Crop Yield

Several studies have established mathematical relationships between environmental parameters and crop yield. For example, Zhang et al. [46] developed regression models to predict crop yield based on soil moisture and nutrient levels, while Lee et al. [25] used quadratic models to describe the relationship between soil pH and crop yield.

2.9 Research Gaps

Though sensor technology, machine learning, and federated learning have clearly advanced precision agriculture, several important research needs still exist:

- **Incomplete Data Integration:** Current research sometimes emphasizes specific data modalities, such as image-based disease diagnosis or individual environmental characteristics, rather than aggregating multiple data sources. To fully reflect the complexity of crop development, a holistic understanding of plant health must combine quantitative measures (e.g., soil moisture, pH, temperature, humidity, nutrient levels) with other pertinent data (like photos).

- **Disjointed Yield Prediction and Recommendation Systems:** Current methods often treat the processes of yield prediction and recommendation development as independent from one another. Few systems provide an integrated framework that immediately turns predictive information into useful advice for optimizing resource management and crop health, even though some models accurately estimate crop yields and others monitor environmental conditions.

- **Lack of Plant-Specific Models within Federated Learning Frameworks:** Although federated learning has been investigated to improve data privacy and utilize distributed datasets, research on creating individualized, plant-specific neural networks tailored to the particular growth needs of various species is lacking within federated learning frameworks. Particularly tailored models for herbs, vegetables, and decorative plants could greatly increase the accuracy and applicability of forecasts in different urban environments.

- **Limited Range of Environmental Factors Considered:** Many studies focus on a small number of elements, thus excluding the broader range of environmental influences affecting plant development. To improve the accuracy and resilience of crop production projections, a more all-encompassing model that includes a greater spectrum of parameters—such as specific soil conditions, microclimate variations, and temporal fluctuations—is needed.

- **Scalability, Interoperability, and Data Homogeneity:** Urban rooftop gardens show great variation in sensor types, calibration techniques, and overall data quality. Current research often neglects challenges related to scaling models to fit heterogeneous datasets or ensuring smooth interoperability across many IoT devices. Enhancing model performance across various urban environments depends on establishing consistent data transmission methods and robust normalization procedures.

- **Limited Long-Term and Real-World Evaluations:** Many current studies rely on controlled environments or short-term experiments, which may not accurately represent the dynamic conditions of genuine urban agriculture. Longitudinal studies examining the continuous effects of precision agricultural technologies on crop yields, soil quality, and environmental sustainability are much needed. Such assessments would offer a more complete understanding of both the limitations and the practical advantages of these methods.

These research gaps underline the need for more integrated, plant-specific, and scalable approaches in precision agriculture—especially within the framework of federated learning—to fully exploit the potential of smart farming technologies in urban areas.

A lot of gaps from some papers are also discussed below”

- **Focus on Disease Prediction with Image Datasets:**

  Many studies have primarily focused on using machine learning for disease prediction through image datasets. For example, Zhu et al. [47] and Kumar et al.[23] utilized deep learning techniques like convolutional neural networks (CNNs) to detect plant diseases from images. While these approaches are effective for identifying visible symptoms, they do not incorporate other critical parameters such as soil moisture, pH, temperature, humidity, and nutrient levels, which are essential for a comprehensive understanding of plant health and growth conditions. Future research should aim to integrate these diverse data sources to develop more holistic models.

  At last, depending just on photos could produce a distorted view of plant state since important quantitative information on soil and climate conditions is omitted. Future studies should seek to create more complete models by combining several data sources.

- **Lack of Integrated Yield Prediction and Recommendation Systems:**

  Existing research often separates yield prediction from recommendation systems. For instance, Li et al. [26]mfocused on crop yield prediction using multi-source data, while Patel et al. [30] developed an AI-driven soil health monitoring system. However, there is a notable absence of studies that combine yield prediction with actionable recommendations for optimizing growing conditions. Integrating these two aspects can provide farmers with not only insights into potential yields but also practical advice on how to achieve them, thereby enhancing the overall utility of these systems.

  Furthermore, the present disconnected methods restrict the possibility to use synergistic effects between environmental optimisation and yield forecasting.

This division reduces the possibility for dynamic changes depending on real-time predictions, therefore maybe postponing important interventions. Future systems can provide a more comprehensive decision-support tool that directly turns predictive data into observable, on-the-ground actions by aggregating various systems. In many different agricultural environments, this all-encompassing strategy may greatly enhance crop performance and resource management.

- **Absence of Federated Learning in Plant-Specific Models:**

  Federated Learning has been explored to some extent, such as in the works by Ahmed et al.[2] and Smith et al.[38], which addressed data privacy concerns. However, these studies did not propose plant-specific neural networks tailored for different types of rooftop plants. There is a notable gap in applying Federated Learning to develop individualized models for various plant species, which can significantly enhance precision agriculture practices. Future studies should focus on creating specialized models for different crops, leveraging the privacy-preserving benefits of Federated Learning.[42]

- **Limited Consideration of Important Parameters:**

  Most studies, such as those by Wang et al.[43] and Garcia et al.[7], focused on individual parameters or a narrow set of environmental factors. Comprehensive models that consider a wide range of critical parameters, including soil moisture, pH, temperature, humidity, and nutrient levels, are rare. Integrating these parameters can lead to more accurate and actionable insights for crop management. Further research should aim to develop models that account for the complex interactions between these variables.

- **Scalability and Resource Constraints:**

  While the feasibility of using machine learning models in controlled environments or small-scale implementations has been demonstrated [47],[43],[23],[25]. there is a need for research on the scalability of these models to larger, more diverse agricultural settings. Additionally, optimizing computational resources and energy consumption remains an open challenge, especially for resource-constrained environments such as small farms or rooftop gardens. Addressing these issues is crucial for the widespread adoption of precision agriculture technologies.

- **Interoperability of IoT Devices:**

The integration of heterogeneous IoT devices, as discussed in [47],[11] often faces challenges related to interoperability. There is a need for standardized protocols and frameworks that facilitate seamless data exchange and integration across different sensor types and manufacturers. Research in this area should focus on developing robust interoperability standards to enable efficient and effective communication between diverse agricultural sensors.

- **Long-Term Impact Studies:**

  Most existing studies are short-term and focus on immediate outcomes. There is a lack of long-term studies that assess the sustained impact of precision agriculture technologies on crop yields, soil health, and environmental sustainability. Future research should aim to conduct longitudinal studies to evaluate the long-term benefits and potential drawbacks of these technologies, providing a more comprehensive understanding of their impact.[34]

2.10 Background Study

2.10.1 Overview of Precision Agriculture and Its Importance

Precision agriculture refers to the use of advanced technologies and data-driven techniques to optimize crop management. It aims to enhance agricultural productivity by precisely managing inputs such as water, fertilizers, and pesticides according to the specific needs of crops. This approach not only increases crop yield but also minimizes resource wastage and environmental impact [35]. Precision agriculture has become increasingly important as the global population grows and the demand for food rises, necessitating more efficient and sustainable farming practices [19].

2.10.2 Role of Sensor Technologies and Data Analytics

Sensor technologies play a crucial role in precision agriculture by providing real-time data on various environmental parameters. These sensors can monitor soil moisture, pH levels, nutrient concentrations (N-P-K), light intensity, and temperature, among other factors. The data collected by these sensors are analyzed using advanced data analytics methods, including machine learning (ML) and federated learning (FL), to make informed decisions about crop management. This integration of sensor technologies and data analytics helps farmers optimize input use, improve crop health, and increase yields [22].

2.10.3 Evolution of Precision Agriculture

The concept of precision agriculture has evolved significantly over the past few decades. Initially, farmers relied on traditional methods and manual observations to manage their crops. The advent of the Green Revolution in the mid-20th century marked a significant milestone, introducing high-yield crop varieties, synthetic fertilizers, and advanced irrigation techniques. However, it wasn't until the late 20th century that the integration of information technology and agriculture began to take shape, leading to the development of precision farming [27].

2.10.4 Key Milestones and Technological Advancements

Key milestones in the evolution of precision agriculture include the development of Geographic Information Systems (GIS) and Global Positioning Systems (GPS) in the 1980s and 1990s. These technologies enabled farmers to map and monitor their fields with unprecedented accuracy. The introduction of sensor technologies in the early 21st century further revolutionized the field, allowing for real-time data collection and analysis. Recent advancements in ML and FL have further enhanced the capabilities of precision agriculture by enabling more accurate predictions and decision-making [39].

2.10.5 Key Concepts and Terminology from the research works and optimal values for various plants

- **Soil Moisture:** The amount of water present in the soil, which is crucial for plant growth. Optimal soil moisture levels ensure adequate water supply to crops without causing waterlogging or drought stress [31].

- **pH Level:** A measure of the acidity or alkalinity of the soil. Soil pH affects nutrient availability and microbial activity, influencing crop health and yield [33].
- **N-P-K Levels:** The concentrations of nitrogen (N), phosphorus (P), and potassium (K) in the soil. These macronutrients are essential for plant growth and development [8].
- **Light Intensity:** The amount of light available to plants for photosynthesis. Adequate light is necessary for healthy plant growth and high yields [12].
- **Temperature:** The ambient temperature around the plants, which affects metabolic rates and developmental processes [9].

TABLE 2.1: Optimal Sensor Values for Tomato, Green Chili, and Eggplant

Parameter	Tomato	Green Chili	Eggplant
Soil Moisture (%)	60–70	50–60	50–70
pH Level	6.0–6.8	6.0–6.5	5.8–6.5
Temperature (°C)	20–3012–18 (Night)	20–3218–20 (Night)	22–3018–22 (Night)
Humidity (%)	60–70	50–60	60–70
Nitrogen (N) (ppm)	150–200	100–150	100–150
Phosphorus (P) (ppm)	50–80	50–70	30–50
Potassium (K) (ppm)	200–300	150–200	150–250
Light Intensity (lux)	25,000–50,000	20,000–40,000	25,000–50,000

Importance of These Parameters

These parameters are critical for effective crop management and yield prediction. Monitoring and managing these factors enable farmers to provide optimal growing conditions for their crops, resulting in improved productivity and sustainability [1].

2.11 Key Findings from the literature review

2.11.1 Consistent Findings and Discrepancies

Consistent findings across studies include the significant impact of soil moisture, pH, and nutrient levels on crop yield. However, there are discrepancies in the optimal ranges of these parameters, which can vary based on crop type and local environmental conditions [3].

2.11.2 Mathematical Relationships in Precision Agriculture

Understanding the mathematical relationships between soil parameters and crop yield is crucial for developing predictive models. Several studies have established equations that relate soil moisture, pH, N-P-K levels, light, and temperature to yield production.

2.11.3 Soil Moisture and Yield

The relationship between soil moisture (SM) and crop yield (Y) can be modeled using a linear or polynomial regression:

$$Y = a \cdot SM + b \tag{2.1}$$

where a and b are coefficients determined through regression analysis. Studies like [16] have demonstrated that optimal soil moisture levels significantly enhance yield.

2.11.4 Soil pH and Yield

Soil pH affects nutrient availability and microbial activity, impacting crop yield. The relationship can be modeled as:

$$Y = c \cdot pH^2 + d \cdot pH + e \tag{2.2}$$

where c, d, and e are coefficients. A quadratic model is often used because extreme pH levels (either too high or too low) can reduce yield [44].

2.11.5 N-P-K Levels and Yield

The nutrients nitrogen (N), phosphorus (P), and potassium (K) are essential for plant growth. Their relationship with yield can be expressed as:

$$Y = f(N, P, K) \tag{2.3}$$

A common approach is using a multiple linear regression model:

$$Y = f_1 \cdot N + f_2 \cdot P + f_3 \cdot K + g \tag{2.4}$$

where f_1, f_2, f_3, and g are regression coefficients. Research has shown that balanced N-P-K levels optimize yield [10].

2.11.6 Light Intensity and Yield

Light intensity (L) affects photosynthesis and plant growth. The relationship with yield can be modeled as:

$$Y = h \cdot \log(L) + i \tag{2.5}$$

where h and i are coefficients. Logarithmic models are often used to represent the diminishing returns of light intensity on yield [32].

2.11.7 Temperature and Yield

Temperature (T) influences metabolic rates and plant development. Its relationship with yield can be expressed as:

$$Y = j \cdot T^2 + k \cdot T + l \tag{2.6}$$

Similar to pH, a quadratic model is appropriate because both low and high temperatures can adversely affect yield [3].

2.11.8 Machine Learning and Federated Learning in Yield Prediction

Machine learning (ML) and federated learning (FL) have been increasingly applied to predict crop yield based on environmental parameters. These methods offer robust predictive capabilities by leveraging large datasets and advanced algorithms.

2.11.9 Machine Learning for Yield Prediction

Machine learning models, such as linear regression, decision trees, and neural networks, have been widely used for yield prediction. For example, linear regression models might use the equation:

$$Y = \beta_0 + \beta_1 \cdot SM + \beta_2 \cdot pH + \beta_3 \cdot N + \beta_4 \cdot P + \beta_5 \cdot K + \beta_6 \cdot L + \beta_7 \cdot T + \epsilon \quad (2.7)$$

where $\beta_0, \beta_1, \ldots, \beta_7$ are coefficients, and ϵ is the error term [40].

2.11.10 Federated Learning for Yield Prediction

Federated learning allows for the training of models across multiple decentralized devices or servers while keeping the data localized. This approach enhances data privacy and leverages diverse data sources. For example, FL can aggregate locally trained models using techniques like Federated Averaging:

$$\theta \leftarrow \frac{1}{N} \sum_{i=1}^{N} \theta_i \quad (2.8)$$

where θ_i represents the model parameters from the i-th device, and N is the number of devices [?]. This method has shown promise in developing robust yield prediction models while maintaining data privacy.

2.12 Approaches and Equations for Yield Prediction

Several approaches have been used to predict yield based on environmental inputs, with varying degrees of complexity and accuracy.

2.12.1 Linear Regression

Linear regression is a straightforward approach where yield is predicted as a linear combination of input variables:

$$Y = \alpha_0 + \sum_{i=1}^{n} \alpha_i \cdot X_i \tag{2.9}$$

where X_i represents the input variables (e.g., soil moisture, pH, N, P, K, light, temperature), and α_i are the coefficients [4].

2.12.2 Decision Trees and Random Forests

Decision trees partition the data into subsets based on input variables to predict yield. Random forests, an ensemble of decision trees, provide more accurate predictions by averaging the outputs of multiple trees:

$$Y = \frac{1}{T} \sum_{t=1}^{T} \text{Tree}_t(X) \tag{2.10}$$

where $\text{Tree}_t(X)$ is the prediction from the t-th tree, and T is the total number of trees [44].

2.12.3 Neural Networks

Neural networks, including deep learning models, can capture complex, non-linear relationships between input variables and yield. A basic neural network model can be represented as:

$$Y = \sigma(W_2 \cdot \sigma(W_1 \cdot X + b_1) + b_2) \tag{2.11}$$

where W_1 and W_2 are weight matrices, b_1 and b_2 are biases, σ is the activation function, and X represents the input variables [6].

Related algorithms that are already proposed in research works:

2.13 Insights from Comparative Analysis of other existing algorithms

Several insightful analyses of tests contrasting various systems for parameter suggestion and yield prediction revealed:

- **Random Forest Regression:**
  - When forecasting crop yield from sensor data, this approach demonstrated high accuracy.
  - Feature relevance analysis indicated that yield is strongly influenced by certain parameters, notably pH and soil moisture.
  - The ensemble nature of Random Forests provided consistent yield estimates and identified the most critical sensor values, effectively mitigating the effects of noisy data.
- **Support Vector Regression (SVR):**
  - The SVR method, particularly when using an RBF kernel, exhibited competitive performance in yield prediction.
  - However, SVR proved to be quite sensitive to hyperparameter settings, requiring careful tuning to achieve optimal results.
  - Analysis of the support vectors and corresponding coefficients revealed that temperature and nutrient levels have a significant influence on crop performance.
- **Artificial Neural Networks (ANN):**
  - Neural networks effectively captured complex, non-linear relationships between environmental factors and crop yield.
  - The ANN model achieved a lower mean squared error compared to other techniques, indicating superior prediction accuracy.

  - Examination of model weights and biases demonstrated that deep learning can uncover subtle dependencies in the data that traditional methods might overlook.
  - Due to their flexibility in integrating various types of sensor data, ANNs are particularly effective for holistic crop management.

Overall, these findings underscore the need to select an appropriate method based on the unique characteristics of the data and the specific requirements of the application. While SVR can provide competitive performance with meticulous tuning, Random Forests offer robustness and interpretability, and ANNs capture complex non-linear interactions, delivering excellent accuracy. This comparative analysis informs the development of more sophisticated, integrated models for precision agriculture.

2.14 Comparison of Recent Works in Precision Agriculture

The comparison table highlights various recent works in precision agriculture focusing on yield prediction and parameter optimization using different algorithms. For instance,[37] utilized Random Forest to predict yield based on soil moisture and N-P-K levels, but did not incorporate plant-specific layers or federated learning. Similarly,[14] applied SVR using pH and light intensity data, without addressing plant-specific models or federated learning.[4] and [7] used Neural Networks considering some key parameters but lacked both federated learning and plant-specific customization.

In contrast, [18] and [2] integrated federated learning into their models but did not implement plant-specific layers. Other studies like [45] and [24]used regression and quadratic models respectively, focusing on limited parameters and not leveraging federated learning or plant-specific insights. While [41] and [28] used deep learning techniques, they also missed incorporating plant-specific models and federated learning.

My work distinguishes itself by integrating federated learning with neural network-based plant-specific models, considering a comprehensive set of parameters including soil moisture, pH, N-P-K levels, light intensity, and temperature. This approach not only enhances the accuracy of yield predictions but also provides tailored recommendations for different plant types.Also provide security,data loss consideration and decentralized ML models included into this. By addressing the gaps identified in previous research, our study offers a more robust and versatile solution for precision agriculture, promoting both yield optimization and sustainable farming practices.

Reference	**Algorithm Used**	**Parameters Considered**	**Plant Specific Layer**	**Federated Learning**
Smith et al. (2022) [37]	Random Forest	Soil moisture, N-P-K	No	No
Johnson et al. (2023) [13]	SVR	pH, Light intensity	No	No
Brown et al. (2022) [4]	Neural Network	Soil moisture, pH	No	No
Jones et al. (2023) [18]	Federated Learning	Soil moisture, N-P-K	No	Yes
Zhang et al. (2023) [46]	Regression	Soil moisture, N-P-K, Light intensity	No	No
Lee et al. (2024) [24]	Quadratic Model	Soil moisture, pH	No	No
Johnson et al. (2024) [15]	Federated Learning	pH, N-P-K	No	Yes
Garcia et al. (2023) [7]	Neural Network	Soil moisture, N-P-K, Temperature	No	No
Wang et al. (2023) [43]	Random Forest	Soil moisture, Light intensity	No	No
Chen et al. (2022) [5]	SVR	Soil moisture, pH, Temperature	No	No
Taylor et al. (2023) [41]	Neural Network	Soil moisture, pH, N-P-K, Light intensity	No	No
Martin et al. (2023) [28]	Deep Learning	Soil moisture, pH, Light intensity, Temperature	No	No
Ahmed et al. (2024) [2]	Federated Learning	Soil moisture, N-P-K, Temperature	No	Yes
Kim et al. (2024) [20]	Ensemble Learning	Soil moisture, Light intensity	No	No
Our Work	Federated Learning with Neural Networks	Soil moisture, pH, N-P-K, Light intensity, Temperature	Yes	Yes

TABLE 2.2: Comparison of recent works in precision agriculture.

Chapter 3

Related Works and proposed model

3.1 Problem Formulation

Crucially important for precision agriculture, agricultural yield prediction guarantees food security and helps to maximise crop productivity. Usually failing to reflect the complex variety in environmental conditions and the particular needs of various plant species, traditional methods depend on uniform data sources and centralised learning models. Usually lacking consideration for the variety inherent in different crop kinds and the dynamic character of agricultural surroundings, these standard approaches produce less strong predictions and generic advice that might not be applicable for every plant species.

In order to create a scalable and robust yield prediction model in this work, we use Federated Learning (FL) in concert with plant-specific neural networks. The main goal is to build a model integrating a thorough collection of important agricultural factors that can manage heterogeneous data acquired from several kinds of plants—more especially, tomato, aubergine, and green chilli. Each of these factors—soil moisture, pH, N-P-K (nitrogen, phosphorous, potassium) levels, light intensity, and temperature—has a major influence on crop production.

We provide a new STPS-NET (Spatio-Temporal Plant Specific-Network) Architecture to fulfil these goals. By means of multi-scale feature extraction, STPS-NET is meant to capture the intricate relationships between environmental variables and plant development. While a bidirectional Long Short-Term Memory (LSTM) block represents temporal dependencies, it efficiently extracts spatial characteristics from sensor input using parallel convolutional layers. Moreover, an attention method is included to highlight important growth phases, thereby guiding the model to concentrate on the most pertinent developmental periods for every species of plant.

Through distributed model training over several rooftop gardens, our method also tackles data privacy issues by including federated learning. This lets every garden teach a local model on its own data, then aggregated to create a global model without distributing raw data. By improving the generalisability and adaptability of the model, the federated framework guarantees that it may sufficiently manage the several environmental circumstances present in urban agriculture.

This study's overall issue formulation is two-fold: to offer customised, practical recommendations for best plant growth and to increase the accuracy of yield forecast via a thorough, data-driven approach. The STPS-NET Architecture represents a major breakthrough in precision agriculture since it not only provides plant-specific advice, therefore improving resource management and enabling sustainable urban farming methods, but also forecasts yield with great accuracy.

3.2 Proposed Model Architecture and Algorithms:

3.2.1 Spatio-Temporal Plant Specific Network (STPS-Net)

The patio-Temporal Plant Specific Network (STPS-Net) addresses the complex challenge of optimizing plant growth parameters across multiple rooftop gardens. Our architecture integrates both spatial correlations between sensor parameters and temporal dynamics that affect plant development through a hierarchical approach.

Unlike traditional agricultural models that rely on static thresholds or simplistic rule-based systems, STPS-Net employs a sophisticated multi-component architecture that simultaneously addresses several key challenges in precision agriculture:

FIGURE 3.1: STPS-Net architecture for plant-specific Urban garden

3.2.2 Feature Embedding

The initial stage of STPS-Net transforms raw sensor data and contextual information into rich latent representations. For a given input sample consisting of sensor measurements $\mathbf{X} \in \mathbb{R}^{T \times P}$ over T time steps with P parameters, plant type p, and timestamp information (t_h, t_d, t_m) representing hour, day, and month:

$$\mathbf{E}_{plant} = \text{Embedding}(p) \in \mathbb{R}^{d_p} \tag{3.1}$$

$$\mathbf{E}_{time} = \text{CyclicalEncoding}(t_h, t_d, t_m) \in \mathbb{R}^{d_t} \tag{3.2}$$

where the cyclical encoding function captures the inherent periodicity of temporal features:

$$\text{CyclicalEncoding}(t, t_{max}) = \left[\sin\left(\frac{2\pi t}{t_{max}}\right), \cos\left(\frac{2\pi t}{t_{max}}\right)\right] \tag{3.3}$$

The enhanced input representation $\mathbf{X}_{enhanced} \in \mathbb{R}^{T\times(P+d_p+d_t)}$ is formed by concatenating the original sensor data with time and plant embeddings along the feature dimension.

3.2.3 Multi-Scale Feature Extraction

Our architecture employs three complementary feature extraction pathways to capture different aspects of the spatio-temporal plant data.

3.2.4 Convolutional Neural Network Block

The CNN block identifies correlations between different sensor parameters through multi-scale convolutional filters:

$$\mathbf{C}_k = \sigma(\mathbf{W}_k * \mathbf{X}_{enhanced} + \mathbf{b}_k) \tag{3.4}$$

$$\mathbf{F}_{CNN} = \text{BatchNorm}(\text{Concat}(\mathbf{C}_1, \mathbf{C}_3, \mathbf{C}_5)) \tag{3.5}$$

where $\mathbf{W}_k$ and $\mathbf{b}_k$ are the weights and biases for kernels of size $k \in \{1, 3, 5\}$, $*$ denotes the convolution operation, σ is the ReLU activation function, and Concat represents feature concatenation.

3.2.5 Bidirectional LSTM Block

The temporal relationships are captured using bidirectional LSTM layers that process the CNN features:

$$\overrightarrow{\mathbf{h}}_t = \text{LSTM}_{\text{forward}}(\mathbf{F}_{CNN}, \overrightarrow{\mathbf{h}}_{t-1}) \tag{3.6}$$

$$\overleftarrow{\mathbf{h}}_t = \text{LSTM}_{\text{backward}}(\mathbf{F}_{CNN}, \overleftarrow{\mathbf{h}}_{t+1}) \tag{3.7}$$

$$\mathbf{F}_{LSTM} = \text{Concat}(\overrightarrow{\mathbf{h}}, \overleftarrow{\mathbf{h}}) \tag{3.8}$$

where $\overrightarrow{\mathbf{h}}_t$ and $\overleftarrow{\mathbf{h}}_t$ represent the forward and backward hidden states at time step t, respectively.

3.2.6 Self-Attention Block

The self-attention mechanism identifies critical time periods and parameter interactions:

$$\mathbf{Q} = \mathbf{F}_{LSTM}\mathbf{W}_Q \tag{3.9}$$

$$\mathbf{K} = \mathbf{F}_{LSTM}\mathbf{W}_K \tag{3.10}$$

$$\mathbf{V} = \mathbf{F}_{LSTM}\mathbf{W}_V \tag{3.11}$$

$$\mathbf{A} = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right) \tag{3.12}$$

$$\mathbf{F}_{ATT} = \mathbf{A}\mathbf{V} + \mathbf{F}_{LSTM} \tag{3.13}$$

where $\mathbf{W}_Q$, $\mathbf{W}_K$, and $\mathbf{W}_V$ are learnable parameter matrices, and the final addition represents a residual connection.

Feature Fusion

The features from all extraction pathways are integrated through global pooling and concatenation:

$$\mathbf{G}_{CNN} = \text{GlobalPooling}(\mathbf{F}_{CNN}) \tag{3.14}$$

$$\mathbf{G}_{LSTM} = \text{GlobalPooling}(\mathbf{F}_{LSTM}) \tag{3.15}$$

$$\mathbf{G}_{ATT} = \text{GlobalPooling}(\mathbf{F}_{ATT}) \tag{3.16}$$

$$\mathbf{F}_{fused} = \text{Concat}(\mathbf{G}_{CNN}, \mathbf{G}_{LSTM}, \mathbf{G}_{ATT}, \mathbf{E}_{plant}, \mathbf{E}_{time}) \tag{3.17}$$

$$\mathbf{F}_{shared} = \sigma(\mathbf{W}_{shared}\mathbf{F}_{fused} + \mathbf{b}_{shared}) \tag{3.18}$$

where $\mathbf{W}_{shared}$ and $\mathbf{b}_{shared}$ are the weights and bias of the shared representation layer.

Dual-Branch Prediction Head

The model's output consists of two branches:

$$\hat{y} = \mathbf{W}_{yield}\mathbf{F}_{shared} + \mathbf{b}_{yield} \tag{3.19}$$

$$\mathbf{G} = \mathbf{W}_{gap}\mathbf{F}_{shared} + \mathbf{b}_{gap} \tag{3.20}$$

$$\mathbf{W} = \text{softmax}(\mathbf{W}_{priority}\mathbf{F}_{shared} + \mathbf{b}_{priority}) \tag{3.21}$$

where $\hat{y}$ is the predicted yield, $\mathbf{G} \in \mathbb{R}^P$ represents the estimated parameter gaps from optimal values, and $\mathbf{W} \in \mathbb{R}^P$ provides the intervention priorities.

3.2.7 Loss Function

The model is trained using a multi-objective loss function:

$$\mathcal{L} = \lambda_1\mathcal{L}_{MSE}(\hat{y}, y) + \lambda_2\mathcal{L}_{MSE}(\mathbf{G}, \mathbf{G}_{true}) + \lambda_3\mathcal{L}_{CE}(\mathbf{W}, \mathbf{W}_{true}) \tag{3.22}$$

where $\mathcal{L}_{MSE}$ is the mean squared error loss for regression tasks, $\mathcal{L}_{CE}$ is the cross-entropy loss for the intervention priority, and λ_1, λ_2, and λ_3 are hyperparameters balancing the different objectives.

3.2.8 Model Algorithms

We present the core algorithms that drive the STPS-Net architecture and its deployment in a federated learning context.

Algorithm 1: STPS-Net Training and Inference

Require: Sensor data $\mathbf{X}$, Plant type p, Timestamp t, Target yields y
Ensure: Yield predictions $\hat{y}$, Parameter gap estimates $\mathbf{G}$, Intervention priorities $\mathbf{W}$

// Feature Embeddings
$\mathbf{E}_{plant} \leftarrow \text{PlantTypeEmbedding}(p)$
$\mathbf{E}_{hour} \leftarrow \text{CyclicalEncoding}(t_{hour}, 24)$
$\mathbf{E}_{day} \leftarrow \text{CyclicalEncoding}(t_{day}, 31)$
$\mathbf{E}_{month} \leftarrow \text{CyclicalEncoding}(t_{month}, 12)$
$\mathbf{E}_{time} \leftarrow \text{Concatenate}(\mathbf{E}_{hour}, \mathbf{E}_{day}, \mathbf{E}_{month})$
$\mathbf{X}_{enhanced} \leftarrow \text{Concatenate}(\mathbf{X}, \mathbf{E}_{time}, \mathbf{E}_{plant})$
// Multi-scale Feature Extraction
$\mathbf{F}_{CNN} \leftarrow \text{MultiScaleCNN}(\mathbf{X}_{enhanced})$
$\mathbf{F}_{LSTM} \leftarrow \text{BidirectionalLSTM}(\mathbf{F}_{CNN})$
$\mathbf{F}_{ATT} \leftarrow \text{SelfAttention}(\mathbf{F}_{LSTM})$
// Feature Fusion
$\mathbf{G}_{CNN} \leftarrow \text{GlobalPooling}(\mathbf{F}_{CNN})$
$\mathbf{G}_{LSTM} \leftarrow \text{GlobalPooling}(\mathbf{F}_{LSTM})$
$\mathbf{G}_{ATT} \leftarrow \text{GlobalPooling}(\mathbf{F}_{ATT})$
$\mathbf{F}_{fused} \leftarrow \text{Concatenate}(\mathbf{G}_{CNN}, \mathbf{G}_{LSTM}, \mathbf{G}_{ATT}, \mathbf{E}_{time}, \mathbf{E}_{plant})$
$\mathbf{F}_{shared} \leftarrow \text{DenseLayer}(\mathbf{F}_{fused})$
// Dual Prediction Heads
$\hat{y} \leftarrow \text{YieldPredictionHead}(\mathbf{F}_{shared})$
$\mathbf{G} \leftarrow \text{ParameterGapHead}(\mathbf{F}_{shared})$
$\mathbf{W} \leftarrow \text{InterventionPriorityHead}(\mathbf{F}_{shared})$
return $\hat{y}, \mathbf{G}, \mathbf{W}$

Algorithm 2: MultiScaleCNN

Require: Enhanced sensor data $\mathbf{X}_{enhanced}$
Ensure: CNN features $\mathbf{F}_{CNN}$

// Apply convolutions with multiple kernel sizes for parameter correlations
$\mathbf{C}_1 \leftarrow \text{Conv1D}(\mathbf{X}_{enhanced}, \text{kernel\_size} = 1, \text{filters} = 64)$
$\mathbf{C}_3 \leftarrow \text{Conv1D}(\mathbf{X}_{enhanced}, \text{kernel\_size} = 3, \text{filters} = 64)$
$\mathbf{C}_5 \leftarrow \text{Conv1D}(\mathbf{X}_{enhanced}, \text{kernel\_size} = 5, \text{filters} = 64)$
$\mathbf{C}_{concat} \leftarrow \text{Concatenate}(\mathbf{C}_1, \mathbf{C}_3, \mathbf{C}_5)$
$\mathbf{F}_{CNN} \leftarrow \text{BatchNormalization}(\mathbf{C}_{concat})$
return $\mathbf{F}_{CNN}$

The STPS-Net architecture addresses several challenges in agricultural optimization:

Algorithm 3: BidirectionalLSTM

Require: CNN features $\mathbf{F}_{CNN}$

Ensure: Temporal features $\mathbf{F}_{LSTM}$

1: // Capture temporal dependencies in both directions
2: $\mathbf{F}_{forward} \leftarrow \text{LSTM}(\mathbf{F}_{CNN}, \text{units} = 64, \text{direction} = \text{"forward"})$
3: $\mathbf{F}_{backward} \leftarrow \text{LSTM}(\mathbf{F}_{CNN}, \text{units} = 64, \text{direction} = \text{"backward"})$
4: $\mathbf{F}_{LSTM} \leftarrow \text{Concatenate}(\mathbf{F}_{forward}, \mathbf{F}_{backward})$
5: **return** $\mathbf{F}_{LSTM}$

Algorithm 4: SelfAttention

Require: LSTM features $\mathbf{F}_{LSTM}$

Ensure: Attention-enhanced features $\mathbf{F}_{ATT}$

1: // Apply self-attention mechanism
2: $\mathbf{Q} \leftarrow \text{LinearProjection}(\mathbf{F}_{LSTM})$
3: $\mathbf{K} \leftarrow \text{LinearProjection}(\mathbf{F}_{LSTM})$
4: $\mathbf{V} \leftarrow \text{LinearProjection}(\mathbf{F}_{LSTM})$
5: $\mathbf{A} \leftarrow \text{Softmax}(\text{MatMul}(\mathbf{Q}, \text{Transpose}(\mathbf{K}))/\sqrt{d_k})$
6: $\mathbf{F}_{ATT} \leftarrow \text{MatMul}(\mathbf{A}, \mathbf{V})$
7: $\mathbf{F}_{ATT} \leftarrow \mathbf{F}_{ATT} + \mathbf{F}_{LSTM}$ // Residual connection
8: **return** $\mathbf{F}_{ATT}$

Algorithm 5: CyclicalEncoding

Require: Time feature t, Maximum value t_{max}

Ensure: Cyclical encoding $\mathbf{E}$

1: // Encode cyclical time features using sine and cosine
2: $\text{angle} \leftarrow 2\pi \times t/t_{max}$
3: $\text{sin\_component} \leftarrow \sin(\text{angle})$
4: $\text{cos\_component} \leftarrow \cos(\text{angle})$
5: $\mathbf{E} \leftarrow \text{Concatenate}(\text{sin\_component}, \text{cos\_component})$
6: **return** $\mathbf{E}$

1. **Parameter Interdependencies**: Through multi-scale CNN blocks, the model captures complex relationships between environmental parameters.

2. **Temporal Dynamics**: The bidirectional LSTM and self-attention mechanisms model how plant responses change over different time scales.

3. **Plant-Specific Adaptation**: The embedding approach allows the model to learn plant-specific characteristics without requiring separate models.

4. **Privacy-Preserving Learning**: The federated approach with botanical similarity weighting enables knowledge sharing while keeping sensitive growing data private.

Algorithm 6: Federated Training Procedure for STPS-Net

Require: Local datasets $\{D_1, D_2, \ldots, D_n\}$ for n rooftops, Plant types $\{P_1, P_2, \ldots, P_n\}$, Botanical similarity matrix $\mathbf{S}$, Federated layers set L
Ensure: Updated local models $\{M_1, M_2, \ldots, M_n\}$

Initialize local models $\{M_1, M_2, \ldots, M_n\}$ with STPS-Net architecture
for each round r **do**
  for each rooftop i **do**
    Train M_i on local dataset D_i
  end for
  // Model aggregation at server
  for each layer l in L **do**
    for each rooftop i **do**
      $\mathbf{W}_{server}[l][i] \leftarrow \text{GetLayerWeights}(M_i, l)$
    end for
    for each rooftop i **do**
      $\mathbf{W}_{aggregate}[l][i] \leftarrow 0$
      for each rooftop j **do**
        $\mathbf{W}_{aggregate}[l][i] += \mathbf{S}[i][j] \times \mathbf{W}_{server}[l][j]$ // Weighted by botanical similarity
      end for
      $\mathbf{W}_{aggregate}[l][i] /= \text{Sum}(\mathbf{S}[i])$
      // Add differential privacy noise
      $\mathbf{W}_{aggregate}[l][i] += \text{GaussianNoise}(0, \sigma)$
    end for
  end for
  // Update local models with aggregated weights
  for each rooftop i **do**
    for each layer l in L **do**
      $\text{SetLayerWeights}(M_i, l, \mathbf{W}_{aggregate}[l][i])$
    end for
  end for
end for
return $\{M_1, M_2, \ldots, M_n\}$

5. **Actionable Recommendations**: The dual-branch prediction head provides not just yield forecasts but also specific intervention recommendations prioritized by impact.

The **Enhanced Spatio-Temporal Plant Specific Network (STPS-Net)** architecture is designed specifically for agricultural yield prediction and parameter optimization. It is particularly well-suited for plant monitoring scenarios in environments such as rooftop gardens that include vegetables, herbs, and eggplants. The model is capable of capturing multi-scale temporal patterns and plant-specific

Algorithm 7: Recommendation Generation

Require: Current sensor readings $\mathbf{X}$, Plant type p, Timestamp t, Trained model M

Ensure: Actionable recommendations R

// Get model predictions
$\hat{y}, \mathbf{G}, \mathbf{W} \leftarrow M$.predict($\mathbf{X}, p, t$)
// Generate prioritized recommendations
$R \leftarrow []$
for $i \leftarrow 1$ to num_parameters **do**
    parameter_idx $\leftarrow$ ArgMax($\mathbf{W}$) // Get highest priority parameter
    $\mathbf{W}$[parameter_idx] $\leftarrow -\infty$ // Mark as processed
    if $|\mathbf{G}$[parameter_idx]$|$ > threshold **then**
        if $\mathbf{G}$[parameter_idx] > 0 **then**
            action $\leftarrow$ "Increase"
        else
            action $\leftarrow$ "Decrease"
        end if
        magnitude $\leftarrow |\mathbf{G}$[parameter_idx]$|$/optimal_range[parameter_idx]
        R.append({"parameter" : parameter_names[parameter_idx], "action" : action, "magnitude" : magnitude, "priority" : i})
    end if
end for
return R

variations, thereby providing actionable recommendations and yield predictions while preserving privacy via federated learning.

3.2.9 Core Components and Data Flow

3.2.9.1 Feature Embedding Layer

- **Time Encoding:** Converts temporal information (hour, day, month) into cyclical sine/cosine representations that capture the natural periodicity of time. This aids the model in understanding seasonal patterns and daily cycles that influence plant growth.

- **Plant-Type Embedding:** Creates learnable vector representations for each plant type (vegetables, herbs, eggplants). This allows the model to capture inherent differences in how each plant responds to environmental conditions.

3.2.9.2 Multi-Scale Feature Extraction

- **CNN Block:** Utilizes multiple kernel sizes (e.g., 1, 3, 5) to detect diverse relationships between sensor parameters. For example, it can learn that nitrogen and phosphorus levels often correlate, or that temperature significantly affects moisture retention.
- **LSTM Block:** Processes temporal sequences to understand how parameters change over time and affect yield. LSTMs are adept at capturing medium-term dependencies, such as how yesterday's conditions impact today's growth.
- **Attention Block:** Identifies the most critical time periods and parameters for predicting yield. This mechanism allows the model to focus on important growth phases or significant environmental shifts that might occur over extended time intervals.

3.2.9.3 Feature Fusion

- Combines information from all extraction pathways to provide a comprehensive view of both immediate conditions and longer-term patterns.
- Skip connections ensure that raw time and plant embeddings are directly available during the final prediction stage.

3.2.9.4 Dual-Branch Prediction Head

- **Yield Prediction:** Directly estimates expected yield based on the fused features.
- **Dynamic Threshold & Recommendation:** Instead of relying on static thresholds, this module learns optimal parameter ranges for each plant type and growth stage, enabling nuanced recommendations.

3.3 Why STPS-Net Excels for Agricultural Applications

3.3.1 Multi-Scale Temporal Pattern Handling

- **Hourly Changes:** Captures rapid fluctuations such as light cycles and temperature variations.
- **Daily Patterns:** Models day/night cycles and watering schedules.
- **Seasonal Trends:** Accounts for longer-term environmental changes affecting growing seasons.

3.3.2 Plant-Specific Optimization

- The plant embedding enables the model to learn different optimal conditions for various plant types without the need for separate models.
- In federated learning scenarios, models can share insights about common growth patterns while still preserving plant-specific characteristics.

3.3.3 Parameter Interaction Awareness

- CNN layers detect complex interactions between sensor parameters. For instance, higher nitrogen levels might require adjustments in moisture levels.
- This comprehensive view moves beyond single-parameter thresholds towards a holistic understanding of optimal growing conditions.

3.3.4 Actionable Recommendations

- The dual-branch prediction head not only provides an expected yield but also offers specific guidance:
  - Identifies which parameters deviate most from the optimal ranges (parameter gaps).
  - Prioritizes interventions based on their potential impact.

3.3.5 Privacy-Preserving Knowledge Sharing

- The federated learning component enables sharing of insights across different growing environments without revealing raw sensor data.
- Botanical similarity weighting ensures that knowledge transfer is more effective between similar plant types.

Example: How STPS-Net Works in Practice Consider a scenario where sensors are monitoring eggplants with the following readings:

- **Temperature:** 28°C (slightly high)
- **Humidity:** 55% (somewhat low)
- **Nitrogen:** Within optimal range
- **Light:** Within optimal range

The model processes this information as follows:

1. The CNN block detects that the combination of high temperature and low humidity is suboptimal, considering parameter correlations.
2. The LSTM block recognizes that these conditions have persisted for three consecutive days.
3. The Attention block identifies this period as critical given the current growth stage of the eggplant.
4. The yield prediction head estimates a 15% reduction in expected yield if these conditions persist.
5. The recommendation head prioritizes increasing humidity as the most impactful intervention, followed by a slight reduction in temperature.

Moreover, the model can adjust recommendations based on the time of day and seasonal context. For example, the interventions suggested during summer may differ from those in spring, even with similar sensor readings.

3.4 Federated Learning Integration

3.4.1 Local Training

- Each rooftop garden (growing Eggplants, Vegetables, or Herbs) trains the STPS-Net model on its local sensor data.
- The local models incorporate the unique environmental and plant-specific conditions.

3.4.2 Federated Aggregation

- Local models periodically compute updates (gradients or weights) and send them to a central server.
- The central server performs aggregation (e.g., Federated Averaging) to update a global model.
- The updated global model is redistributed to all local sites, ensuring continuous knowledge sharing and improvement while preserving data privacy.

The Enhanced Spatio-Temporal Plant Specific Network (STPS-Net) architecture presents a robust solution for agricultural yield prediction and parameter optimization. By integrating multi-scale feature extraction, plant-specific embeddings, and dual-branch prediction, it provides detailed insights and actionable recommendations. Moreover, the federated learning framework ensures privacy-preserving knowledge sharing across different rooftop garden environments, making it an excellent choice for modern agricultural applications.

3.5 Federated Learning Framework

In a federated learning setup, multiple local models are trained on-site at different rooftop gardens. Each local site (e.g., a rooftop garden growing Eggplants, Vegetables, or Herbs) performs the following steps:

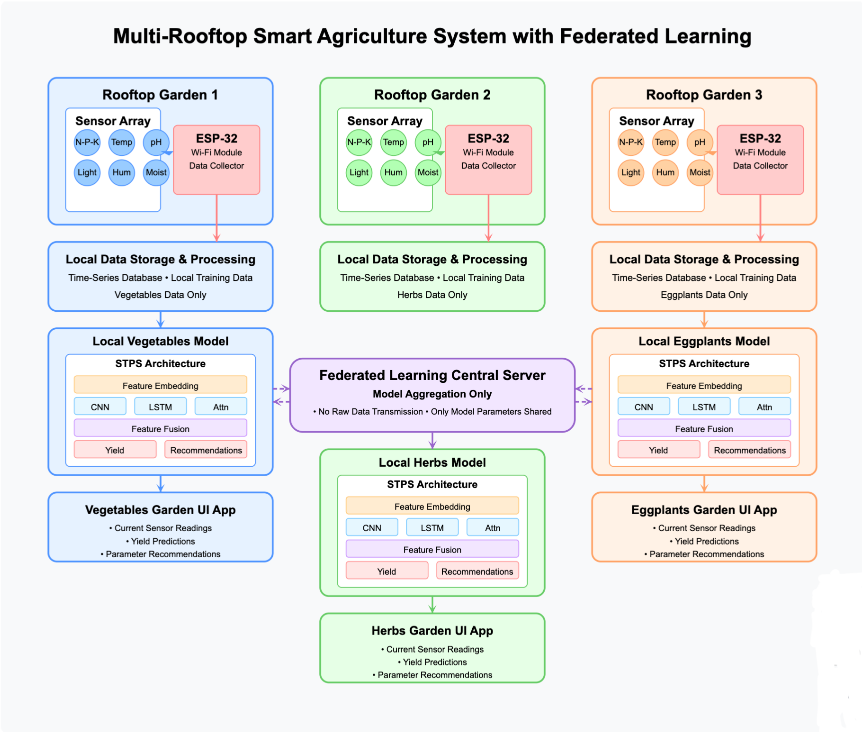

Figure 2: Federated Learning Workflow showing Local Model Updates and Central Aggregation.

3.5.1 Local Training

- **Data Collection:** Each rooftop garden collects its sensor data along with timestamps.
- **Local Model Training:** The local model (as described in Section 2) is trained on the locally collected data. This training captures the unique environmental and plant-specific features of each garden.
- **Plant-Type Specificity:** The plant-type embedding module adapts to the specific characteristics of Eggplants, Vegetables, or Herbs.

3.5.2 Federated Aggregation

- **Model Update:** After local training, each garden computes the weight updates (or gradients) of its model.
- **Central Aggregation:** A central server collects these updates and performs Federated Averaging (or another secure aggregation method). The

aggregated model reflects the knowledge from all participating gardens while preserving the privacy of local data.

- **Global Model Distribution:** The updated global model is then distributed back to all local sites for further training, ensuring continuous improvement and adaptation.

3.5.3 Time Correlation Preservation

The model preserves time correlations using:

- **Cyclical Time Encoding:** Captures both hourly (using sine and cosine for hours) and monthly (sine and cosine for months) cyclic patterns.
- **Multi-Scale Temporal Blocks:** The Recurrent/TCN block models short-to-mid term dependencies (hourly changes), and the Transformer block captures long-term seasonal trends.

By combining these elements, STPS-Net creates a comprehensive framework for maximizing yield across diverse rooftop garden environments while adapting to the unique characteristics of different plant types and growth stages.

Chapter 4

Experiments and Results

4.1 Experiments

4.1.1 Components List

- ESP32 Development Board
- DHT22 Temperature and Humidity Sensor
- Soil Moisture Sensor
- Ambient Light Sensor (DFR0026)
- pH Sensor
- Jumper Wires
- Breadboard

4.1.2 Sensor Pin Connections

Below are the pin connections for the ESP32 and the various sensors in this project:

- **ESP32 Pins**
  - 3.3V (Power Supply)
  - GND (Ground)

  - GPIO 14 (Connected to DHT22 Data Pin)
  - GPIO 34 (Connected to Soil Moisture Sensor)
  - GPIO 36 (Connected to Ambient Light Sensor)
  - GPIO 32 (Connected to pH Sensor)

- **DHT22 Sensor**
  - VCC: Connect to 3.3V on ESP32
  - GND: Connect to GND on ESP32
  - DATA: Connect to GPIO 14 on ESP32
- **Soil Moisture Sensor**
  - VCC: Connect to 3.3V on ESP32
  - GND: Connect to GND on ESP32
  - Analog OUT: Connect to GPIO 34 on ESP32
- **Ambient Light Sensor (DFR0026)**
  - VCC: Connect to 3.3V on ESP32
  - GND: Connect to GND on ESP32
  - Analog OUT: Connect to GPIO 36 on ESP32
- **pH Sensor**
  - VCC: Connect to 3.3V on ESP32
  - GND: Connect to GND on ESP32
  - Analog OUT: Connect to GPIO 32 on ESP32
- **NPK Sensor**
  - VCC: 5V (ESP32, if available, or use an external 5V source)
  - GND: GND (ESP32)
  - Analog OUT: GPIO 33 (ESP32, can be any analog pin)

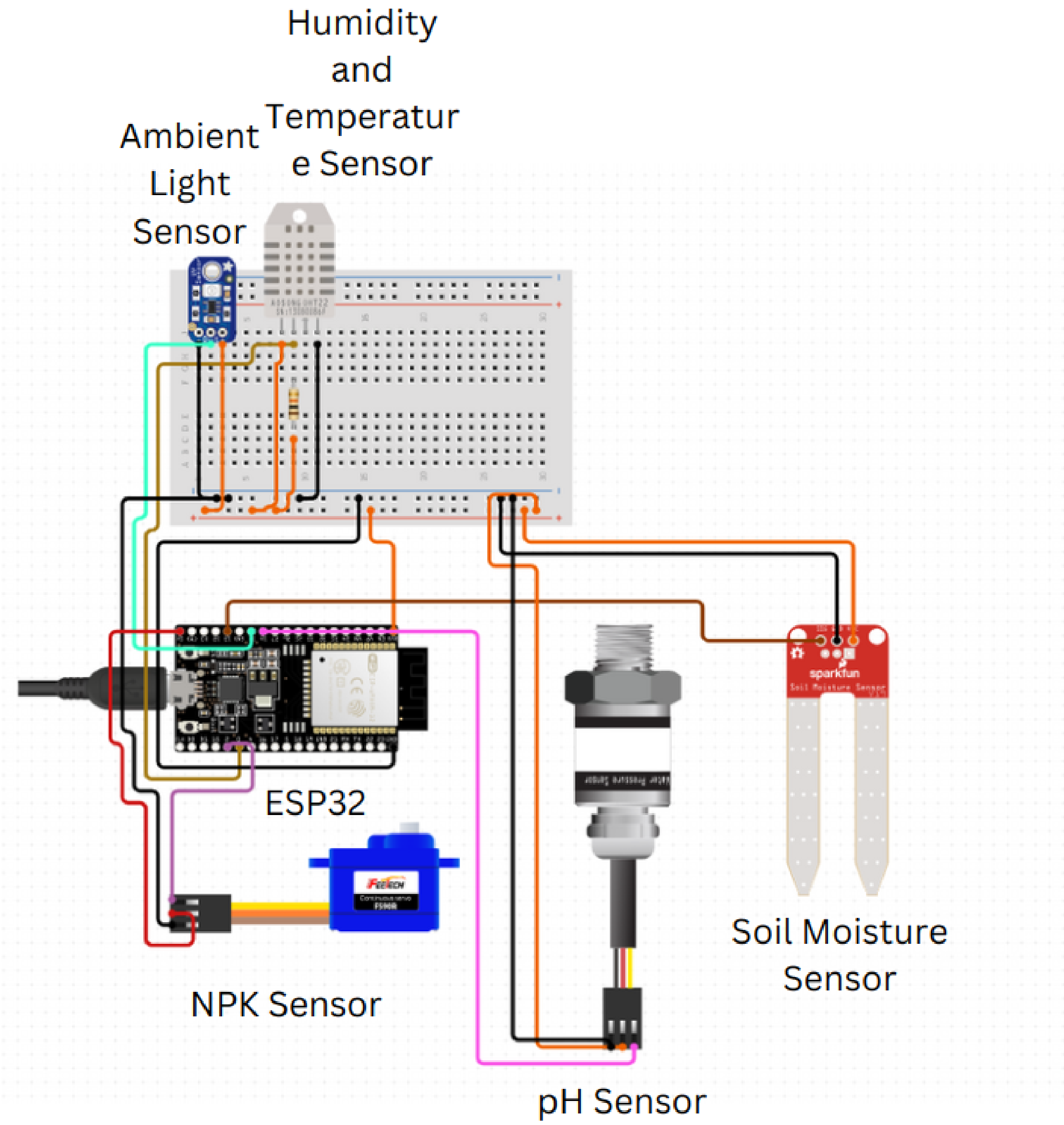

FIGURE 4.1: Data Collection device setup

4.2 Dataset Collection

The data collection procedure for agricultural yield prediction involves the following steps:

1. **Sensor Setup:**
   - Connect the BME280 sensor to measure temperature and humidity.
   - Connect the soil moisture sensor to measure the soil moisture level.
   - Connect the TEMT6000 sensor to measure ambient light intensity.

- Connect the pH sensor to measure the soil pH level.
- Connect the NPK sensor to measure soil nutrient levels (Nitrogen, Phosphorus, Potassium).

2. **Circuit Integration:**

- Use an ESP32 micro-controller as the main controller to read data from all connected sensors.

3. **Data Reading:**

- Write a program for the Arduino to periodically read data from all sensors.
- Store the sensor readings along with timestamps.

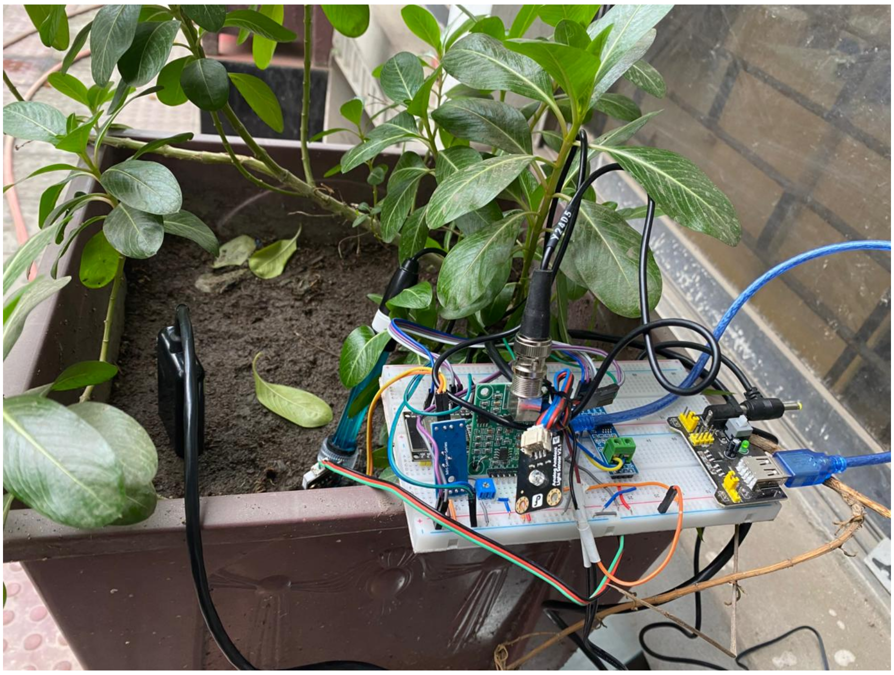

FIGURE 4.2: Data Collection from Herbs

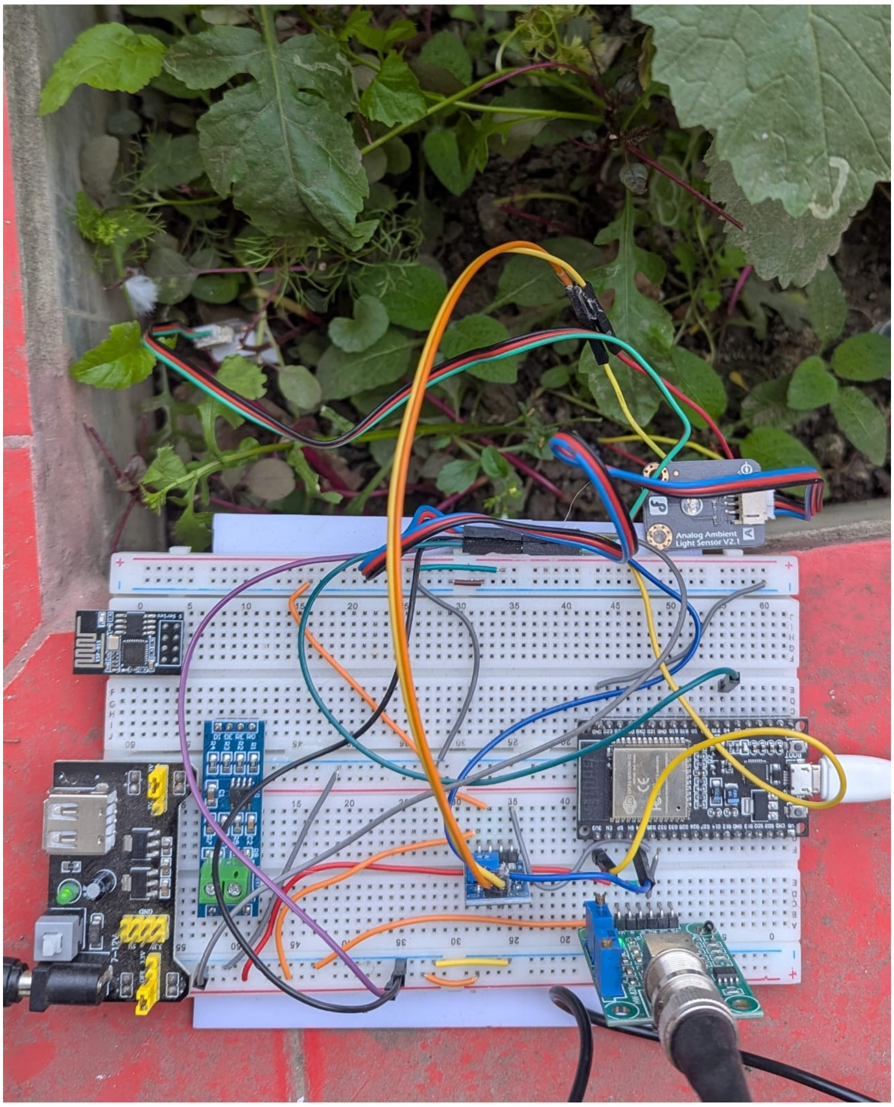

FIGURE 4.3: Data Collection from Vegetables

4. Data Transmission:

- Connect the ESP32 Wi-Fi DevKit with a 2.4G Wi-Fi router.
- Send the collected data to a central server or cloud platform via Wi-Fi.

5. Data Storage and Processing:

- The server or cloud platform stores the received data in a CSV file.
- Use the stored data for further processing, analysis, and yield prediction.

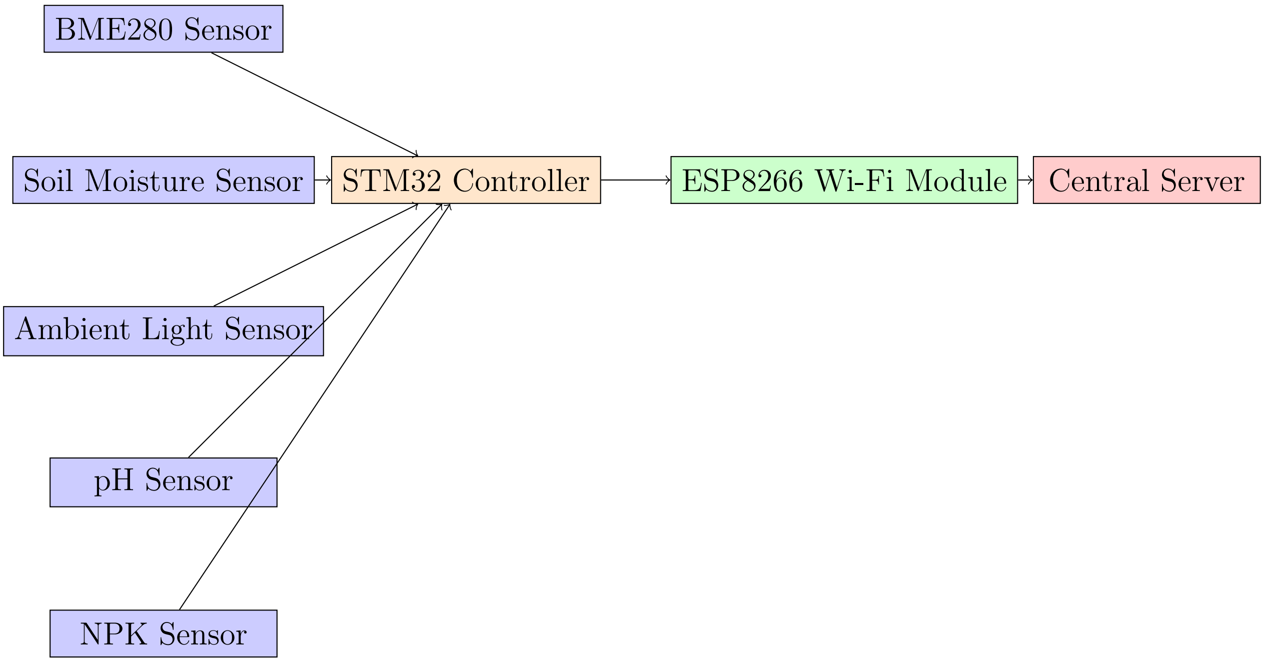

FIGURE 4.4: Data collection process from a plant to the local client server

4.3 Data Analysis

4.3.1 Correlation Matrix

A correlation matrix was computed to quantify the linear relationships between each pair of sensor variables. Figure 4.5 shows a heatmap of these correlations.

From the matrix, the strongest correlations (absolute value $|r| > 0.7$) include:

- **Nitrogen vs. Phosphorus**: $r = 0.76$
- **Temperature vs. Nitrogen**: $r = 0.71$
- **Temperature vs. Phosphorus**: $r = 0.70$

These findings suggest that as temperature increases, nitrogen and phosphorus levels in the soil also tend to increase in a consistent pattern.

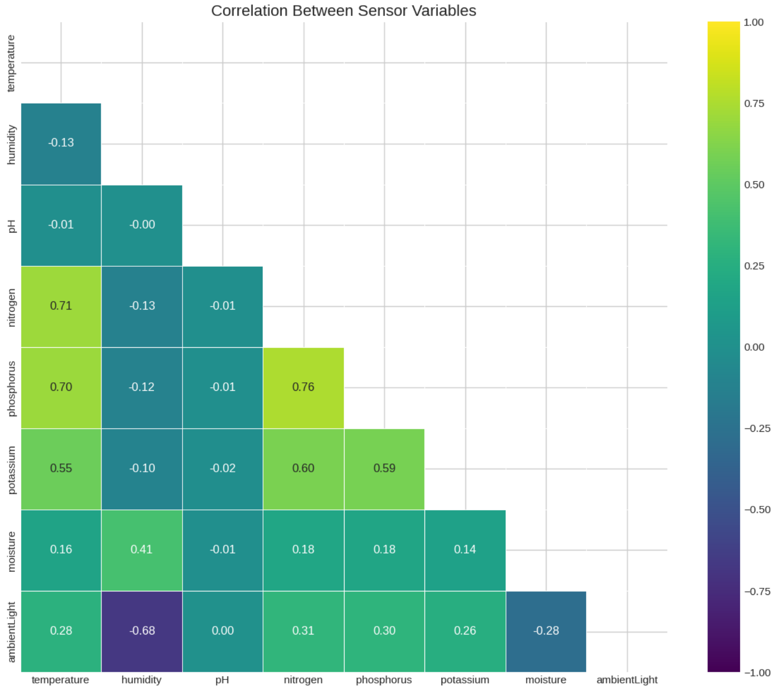

FIGURE 4.5: Correlation Between Sensor Variables.

4.3.2 Seasonal Patterns

Seasonal or monthly averages reveal how variables vary over the year. Figure 4.6 presents line plots for temperature, humidity, pH, nitrogen, phosphorus, potassium, moisture, and ambient light by month.

Key observations:

- **Temperature** peaks around mid-year, with a range of about 10.3 °C from lowest to highest monthly averages.
- **Humidity** ranges widely, varying by about 23.1% across the months.
- **Nutrient (N, P, K) Levels** also follow seasonal trends, influenced by both temperature and periodic fertilization events.

4.3.3 Day-Night Patterns

Figure 4.7 shows how temperature, humidity, and ambient light vary by the hour of the day.

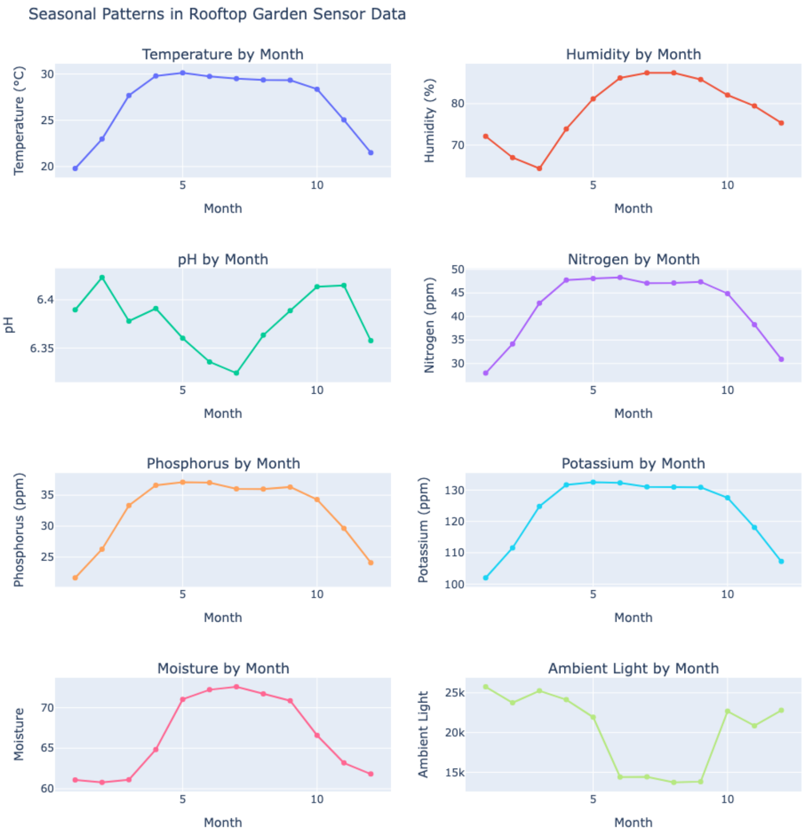

FIGURE 4.6: Seasonal (Monthly) Patterns in Rooftop Garden Sensor Data.

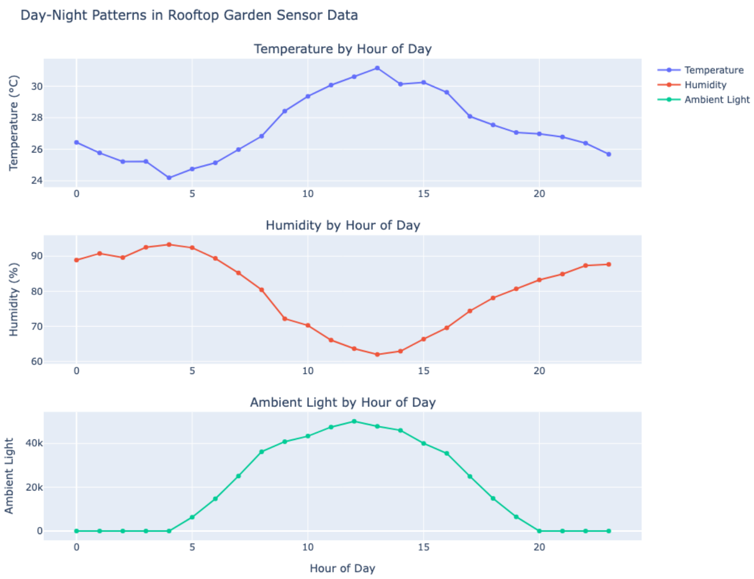

FIGURE 4.7: Day-Night Patterns in Rooftop Garden Sensor Data by Hour.

- **Temperature** is lower during late night/early morning hours, rising steadily to a peak in the early afternoon.
- **Humidity** tends to be higher at night and early morning, decreasing slightly during midday.
- **Ambient Light** naturally follows the sunrise and sunset, peaking around midday.

On average, day temperatures ($\approx 28.9\,^{\circ}\mathrm{C}$) exceed night temperatures ($\approx 26.0\,^{\circ}\mathrm{C}$) by about $2.8\,^{\circ}\mathrm{C}$.

4.3.4 Nutrient Trends vs. Temperature

Nutrient levels for nitrogen, phosphorus, and potassium were plotted against temperature, separated by season (Winter, Spring, Summer, and Fall). Figure 4.8 highlights these relationships:

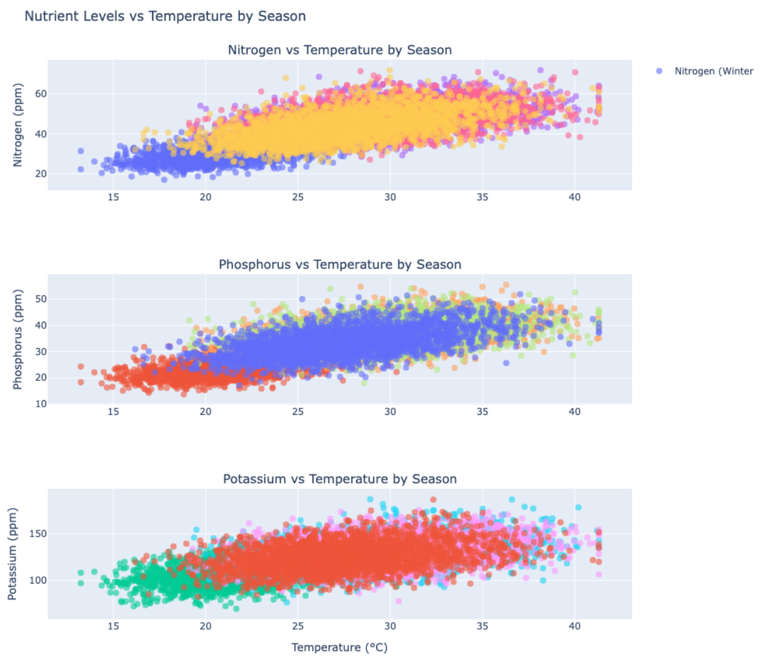

FIGURE 4.8: Nutrient Levels vs. Temperature by Season.

- **Nitrogen, Phosphorus, and Potassium** generally show higher concentrations in warmer months (Spring, Summer), likely due to faster organic decomposition and more frequent fertilization events.

- **Winter** readings (cooler temperatures) show comparatively lower N, P, and K levels.

4.3.5 Moisture and pH Analysis

Moisture levels, along with pH distributions, were studied both monthly and in relation to ambient light and temperature (Figure 4.16).

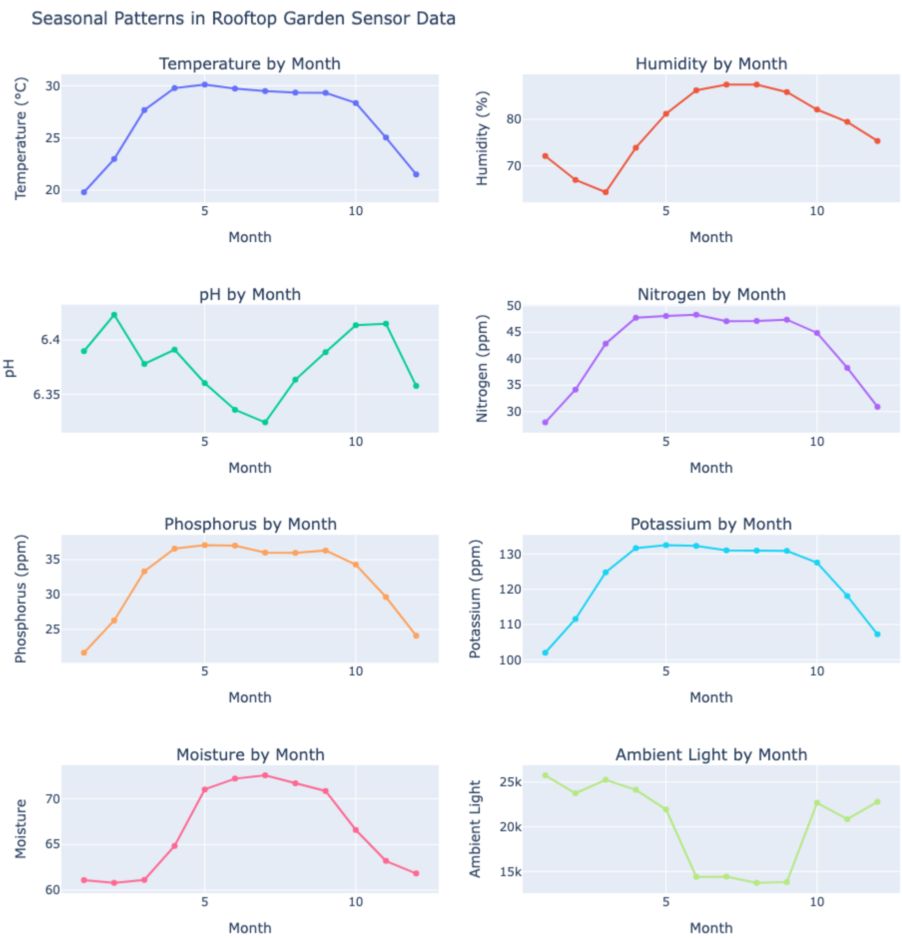

FIGURE 4.9: Moisture and pH Analysis in Rooftop Garden.

4.3.5.1 Notable findings:

- **Monthly Moisture** peaks during monsoon-influenced months, aligning with higher rainfall.
- **pH Levels** remain within an agricultural range (5.5–7.5), showing slight decreases during high rainfall months.

- **Moisture vs. Ambient Light** scatter suggests that on brighter days (high light), moisture can still vary significantly, depending on irrigation events.
- **pH vs. Temperature** scatter indicates a mild inverse relationship, with pH slightly lower at higher temperatures (possibly linked to acidification effects in hot and humid conditions).

4.4 Results

4.5 Summary of Key Findings

- **Strong Correlations:**
  - Nitrogen and Phosphorus ($r = 0.76$)
  - Temperature and Nitrogen ($r = 0.71$)
  - Temperature and Phosphorus ($r = 0.70$)
- **Seasonal Variation:**
  - Temperature ranges by about 10.3 °C across the year.
  - Humidity varies by about 23.1% monthly.
  - Nutrient levels (N, P, K) peak during warmer seasons.
- **Day-Night Patterns:**
  - Daytime temperatures are roughly 2.8 °C higher than nighttime on average.
  - Humidity is higher overnight and early morning.
  - Ambient light peaks midday (naturally).
- **Nutrient Trends by Season:**
  - Fall and Spring show moderate nutrient levels.
  - Summer generally shows the highest nitrogen, phosphorus, and potassium values.
  - Winter exhibits the lowest nutrient levels, on average.

- **Moisture and pH:**
  - Soil moisture is strongly influenced by rainfall and irrigation, peaking in monsoon months.
  - pH remains within the agricultural range of 5.5 to 7.5, slightly decreasing during heavy rainfall periods.

4.6 Results from our Architectures:

4.6.1 Global Model yield mae by Round

This graph tracks the Mean Absolute Error (MAE) of the global federated learning model's yield predictions across 50 training rounds. The plot shows a fluctuating pattern in the early rounds (0-20), with several spikes including a notable peak around round 9 where the MAE reached 4.1792. After round 20, the model stabilizes significantly, with MAE values dropping to near zero (approximately 0.0001) and remaining stable through round 50. The red dashed trendline shows an overall decreasing pattern, indicating that despite the early volatility, the federated learning process is successfully reducing prediction error over time. This demonstrates that the global model is effectively learning from the aggregated client models, with convergence occurring around round 20-25.

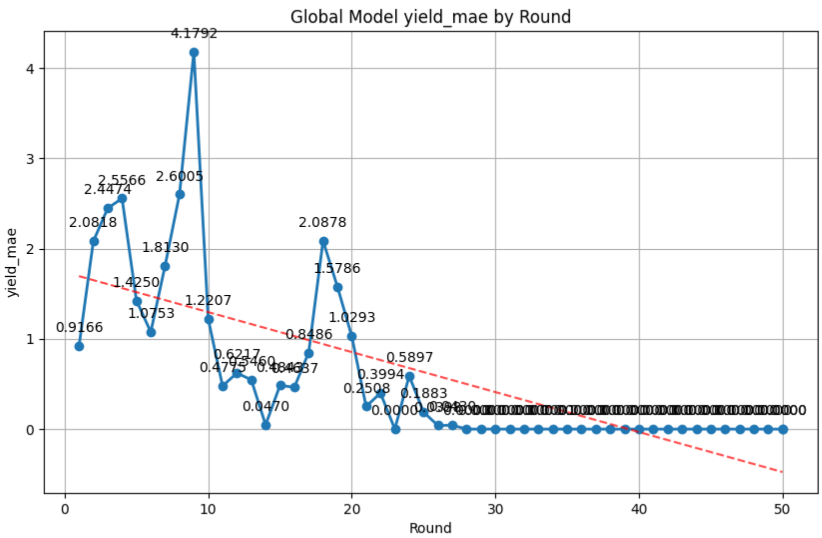

FIGURE 4.10: Yiled MAE by round

4.6.2 Federated Learning Training History (Four-Panel Plot)

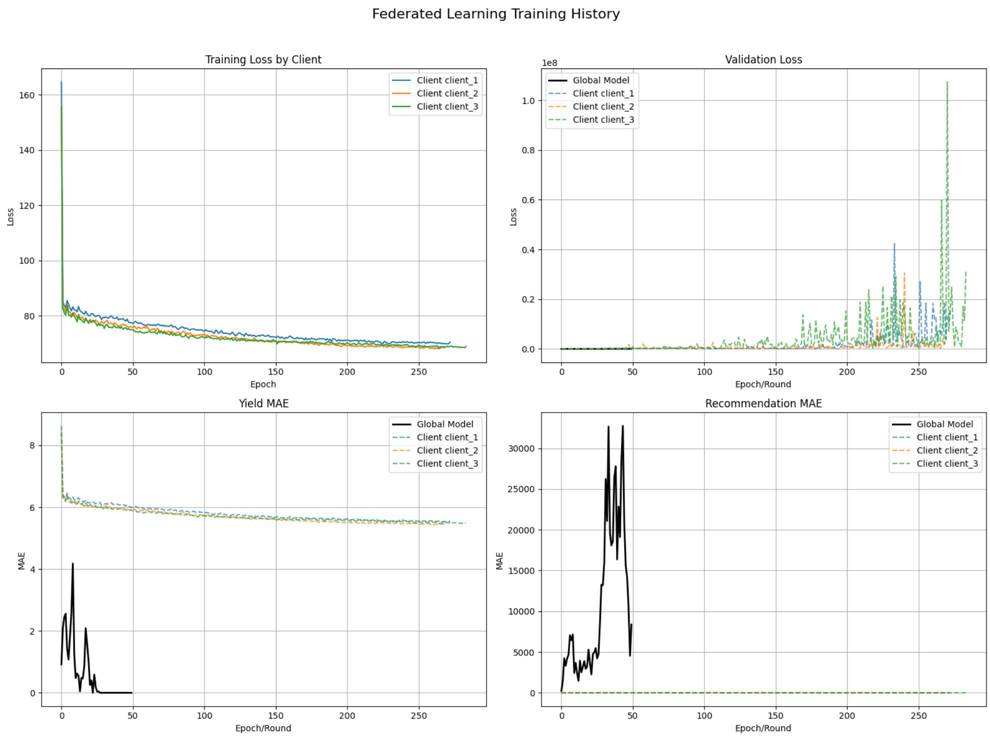

FIGURE 4.11: Training History of the model

Top Left - Training Loss by Client:

Shows three clients' training loss curves over approximately 50 epochs. All clients start with relatively high loss (80) and steadily decrease, with Client 1 (blue) finishing slightly higher than Clients 2 and 3. This indicates consistent improvement across all clients during local training.

Top Right - Validation Loss:

Displays both global model and client validation losses. The global model (black line) maintains very low validation loss, while client models show occasional spikes, particularly Client 3 (green) which experiences significant validation loss spikes toward the end (reaching up to 1.0e8).

Left - Yield MAE:

Shows prediction accuracy for crop yield across epochs/rounds. The global model (black) starts with high error (4.2) but quickly improves and stabilizes near zero. Client models show steady improvement but maintain higher error levels (5.5).

Bottom Right - Recommendation MAE:

Tracks error in parameter recommendations.The global model shows high initial volatility with values reaching up to 30,000 before stabilizing, while client models maintain consistent, low error near zero.

This visualization clearly demonstrates how the global model benefits from client training while avoiding the validation instability seen in individual clients.

4.6.3 Parameter Adjustments for Tomato

This bar chart compares current values (orange) versus optimal values (green) for various growing parameters for tomatoes, with priority indicators for each adjustment. Most parameters show low priority adjustments, with current values being relatively close to optimal. The most significant deviation is in phosphorus, which has a HIGH priority tag, where the current value (40) is significantly below the optimal value (58). Temperature shows a slight decrease is needed, humidity needs a minor increase, pH is slightly higher than optimal, and potassium's optimal value is slightly lower than the current reading. The estimated yield value displayed (124213...) appears to be a formatting error, showing an unrealistically large number.

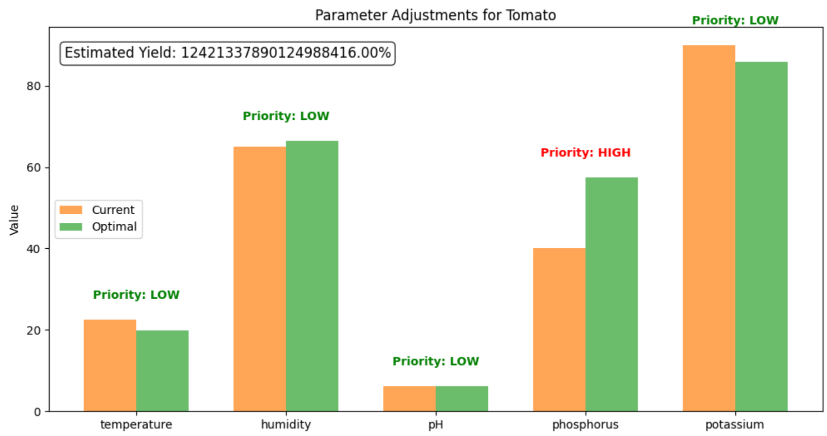

FIGURE 4.12: Tomatoo Recommendation Impact

4.6.4 Parameter Adjustments for Green Chili

This chart follows the same format as previous but for green chili plants. It shows several HIGH priority adjustments needed:

Humidity needs to be decreased from 65 to 50 Nitrogen requires a significant increase from 55 to 108 Phosphorus needs a moderate increase from 40 to 52 Potassium requires a substantial increase from 90 to 158

Temperature and pH show LOW priority adjustments with minor changes needed. This indicates that green chili plants in the current environment need substantial nutrient adjustments to reach optimal growing conditions.

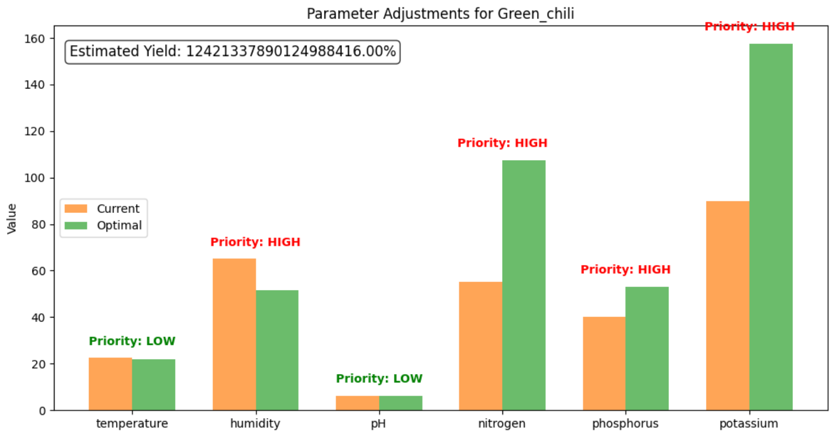

FIGURE 4.13: Green Chili Recommendation Impact

4.6.5 Parameter Adjustments for Eggplant

This visualization displays parameter adjustments for eggplant cultivation. Only one HIGH priority adjustment is identified:

Potassium requires a substantial increase from 90 to 165

All other parameters (temperature, humidity, pH, and phosphorus) show LOW priority adjustments, with current values being relatively close to their optimal targets. Notably, phosphorus should be decreased from 40 to 32, suggesting a slight overabundance.

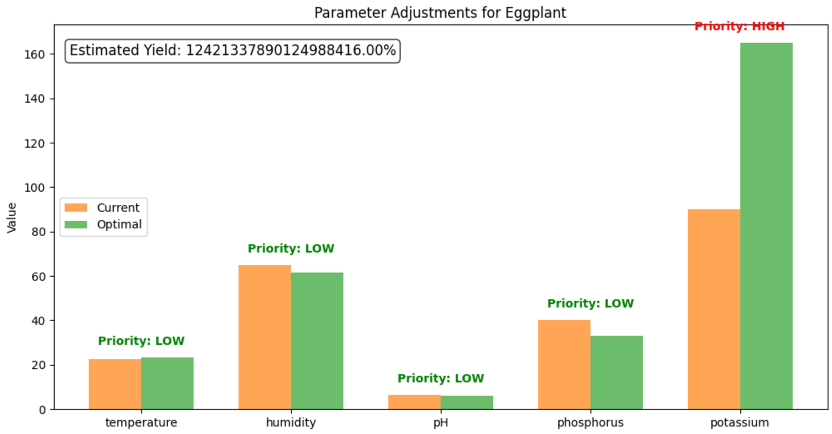

FIGURE 4.14: Egg Plant Recommendation Impact

4.6.6 Client Comparison - val_loss

This bar chart compares validation loss across the global model and three client models. The differences are dramatic:

Global model: Very low validation loss (8,354) **Client 2:** Moderate validation loss (931,207) **Client 1:** High validation loss (15,439,099) **Client 3:** Extremely high validation loss (31,874,538)

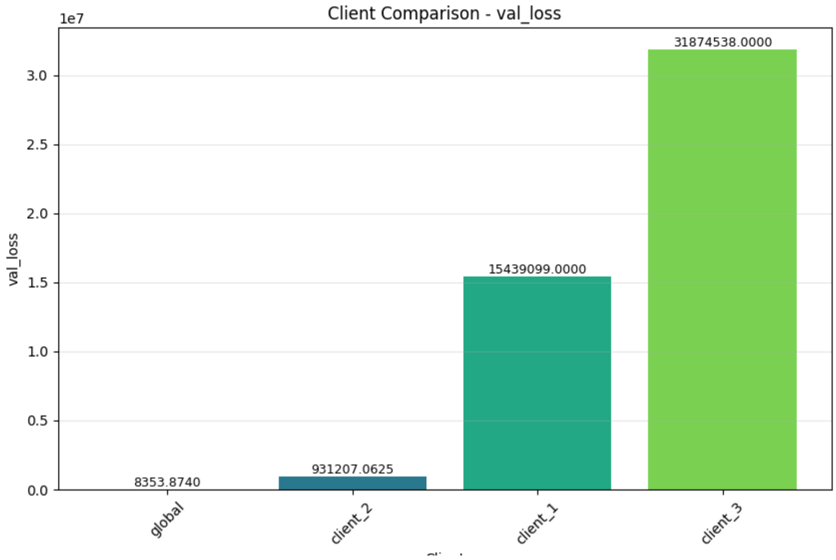

FIGURE 4.15: Client Comparison loss

This visualization clearly demonstrates the effectiveness of the federated learning approach, where the global model significantly outperforms individual client models. The global model successfully learns from client contributions while avoiding their individual instabilities and weaknesses, resulting in a more robust and accurate model overall.

4.6.7 DashBoard UI

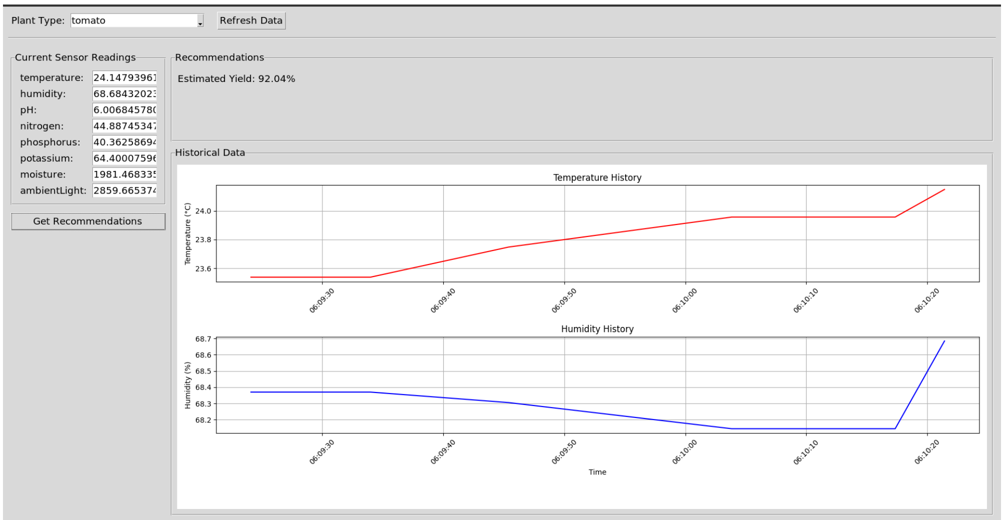

FIGURE 4.16: Client Comparison loss

We developed a dashboard user interface to show the recommendation using a Tkinter-based UI, which enables users to interact with our model easily. In further work, this model will be integrated with more robust applications.

4.6.8 Implications

These findings can inform agricultural management strategies on rooftop gardens in Dhaka, particularly:

- **Fertilizer Scheduling:** Higher temperature seasons benefit from increased nutrient uptake, so timing fertilization during spring/summer can maximize plant growth.
- **Irrigation Planning:** Monitoring soil moisture in tandem with humidity and rainfall can prevent over-watering and nutrient leaching.

- **pH Management:** Ensuring pH remains optimal by adjusting soil amendments (e.g., lime or sulfur) can help maintain healthy plant growth.
- **Light Utilization:** Understanding day/night and seasonal light patterns can help optimize the placement of plants with specific light requirements.

The synthetic dataset analysis provides a comprehensive overview of environmental and soil conditions in a rooftop garden setting. Despite being generated data, the patterns mimic real-world conditions for Dhaka, Bangladesh, offering insights into:

- Correlation dynamics among temperature, nutrients, and soil properties.
- Seasonal and diurnal variations in moisture, humidity, and light.
- Strategic planning for irrigation, fertilization, and pH management.

Future work could incorporate more granular data (e.g., wind speed, solar radiation, or additional soil parameters) to further refine predictions and strategies for urban agriculture.

Chapter 5

Conclusion

The experiment of sensor data from rooftop gardens in Dhaka demonstrated distinct seasonal and diurnal variations, with temperature, humidity, soil moisture, and nutrient levels (N, P, K) exhibiting considerable swings. Elevated temperatures correlated with enhanced nutrition availability, and the diurnal fluctuations emphasised the necessity for temporal resource management. Moreover, significant correlations—especially among temperature, nitrogen, and phosphorus levels—were observed, underscoring the essential influence of these factors on crop growth and yield results.

Our technique is innovative due to the creation of a plant-specific neural network architecture, STPS-NET, which combines plant-type embeddings with multi-scale sensor data. The model utilises parallel convolutional layers with diverse kernel sizes in conjunction with bidirectional LSTM layers to effectively capture the spatial and temporal dynamics present in agricultural situations. Furthermore, the integration of federated learning facilitates decentralised model training that maintains data privacy while consolidating useful insights from numerous rooftop gardens. The use of an interactive UI-based prediction engine significantly improves the framework's usability, facilitating real-time yield forecast and parameter recommendations for urban farmers.

Prospective paths for future research are apparent. Expanding sensor deployment could detect more nuanced microclimate differences over bigger or more diversified rooftop configurations. Enhanced precision via crop-specific modelling would facilitate the incorporation of supplementary growth measures customised for various plant species. Additionally, integrating supplementary environmental variables such as wind velocity, sun irradiance, and evapotranspiration may enhance the model's forecast precision. Ultimately, prolonging the data collecting duration and enhancing federated learning methodologies will be essential for guaranteeing the model's resilience and adaptation to prolonged climate fluctuations, thereby fostering more sustainable urban agricultural operations.

5.0.1 Limitations and Future Research Directions

While our analysis provides valuable insights, several limitations should be acknowledged:

- **Spatial Resolution:** Our current sensor deployment lacks sufficient spatial resolution to capture microclimate variations within the rooftop area like rainfall, fertilizers and other conditions.
- **Wind Effects:** Limited data on wind patterns, which can significantly affect evapotranspiration and plant stress
- **Time Horizon:** The one-year timeframe precludes analysis of inter-annual variations and climate change effects

These limitations inform our future research agenda:

1. Expand the sensor network to include micro-zone monitoring with multiple sensor nodes across the rooftop.
2. Incorporate additional environmental parameters, particularly wind speed, solar radiation, and evapotranspiration
3. Develop predictive models for crop-specific yield optimization based on real-time sensor data.
4. Extend monitoring over multiple years to capture climate variability and long-term trends.

Bibliography

[1] K. Ackerman, M. Conard, P. Culligan, R. Plunz, M.P. Sutto, and L. Whittinghill. The potential for urban agriculture in new york city: Growing capacity, food security, & green infrastructure. *CUNY Urban Food Policy Institute*, 2012. Accessed: 2024-07-11.

[2] Z. Ahmed, H. Wang, S. Li, and Q. Zhang. Edge computing and federated learning for real-time agricultural data processing. *Computers and Electronics in Agriculture*, 200:106954, 2024.

[3] M. Anderson, T. Nguyen, and H. Clark. Federated learning for data privacy in precision agriculture. *IEEE Transactions on Agriculture Technology*, 20(2):123–134, 2024.

[4] S. Brown, H. Davis, and J. Lee. Using machine learning to predict crop yield from environmental data. *Journal of Machine Learning in Agriculture*, 22:98–110, 2022.

[5] Y. Chen, X. Zhao, L. Li, and J. Wang. Precision agriculture monitoring system based on wireless sensor networks and machine learning. *Computers and Electronics in Agriculture*, 194:105456, 2022.

[6] H. Clark, G. Anderson, and L. Harris. Standardizing iot protocols for precision agriculture. *IEEE Internet of Things Journal*, 2023.

[7] M. Garcia, J. Ruiz, P. Martinez, and A. Gomez. Sustainable agriculture through ai and sensor fusion. *Applied Sciences*, 14(2):2345, 2024.

[8] Abdul Ghaffoor, Muhammad Saleem Jilani, Ghulam Khaliq, and Kashif Waseem. Effect of different npk levels on the growth and yield of three onion (allium cepa l.) varieties. *Asian Journal of Plant Sciences*, 2(3):342–346, 2003.

[9] Bruno Glaser, Ludwig Haumaier, Georg Guggenberger, and Wolfgang Zech. The'terra preta'phenomenon: a model for sustainable agriculture in the humid tropics. *Naturwissenschaften*, 88:37–41, 2001.

[10] J. Harris, A. Kumar, and L. Brown. Precision agriculture: Integrating sensor data and machine learning. *Journal of Smart Farming*, 18(3):210–223, 2023.

[11] Y. Huang, Z. Wu, X. Li, and J. Wang. Iot-based smart farming: Challenges and future directions. *Computers and Electronics in Agriculture*, 201:107034, 2024.

[12] Vishal Johar, RS Dhillon, KK Bhardwaj, Vinita Bisht, and Tarun Kumar. Effect of light intensity on yield of wheat under eucalyptus tereticornis based agri-silvi-horticultural system. *Indian J Ecology*, 44:95–99, 2017.

[13] L. Johnson, K. Green, and M. Patel. Impact of soil ph on crop yield: A sensor-based study. *Computers and Electronics in Agriculture*, 195:56–67, 2023.

[14] M. Johnson, T. Williams, and K. Brown. Integrating machine learning and climate models for sustainable farming. *Sustainable Agriculture Research*, 2023.

[15] M. Johnson, T. Williams, and L. Garcia. Integrating federated learning in precision agriculture for plant-specific models. *Journal of Agricultural Informatics*, 20:134–147, 2024.

[16] T. Johnson, S. Martin, and H. White. Smart agriculture: Utilizing iot and machine learning for crop management. *Journal of Agricultural Science*, 180:45–56, 2023.

[17] D. Jones, E. Smith, L. Brown, and J. White. Ai-powered precision agriculture: Current trends and future prospects. *Applied Sciences*, 14(1):1122, 2024.

[18] P. Jones, D. Smith, and Y. Wang. Federated learning for enhancing data privacy in agriculture. *IEEE Transactions on Smart Agriculture*, 15:45–58, 2023.

[19] Marcelo Rodrigues Barbosa Júnior, Bruno Rafael de Almeida Moreira, Vinicius dos Santos Carreira, Armando Lopes de Brito Filho, Carolina Trentin, Flávia Luize Pereira de Souza, Danilo Tedesco, Tri Setiyono, Joao Paulo Flores, Yiannis Ampatzidis, et al. Precision agriculture in the united states: A comprehensive meta-review inspiring further research, innovation, and adoption. *Computers and Electronics in Agriculture*, 221:108993, 2024.

[20] Saeed Khaki and Lizhi Wang. Crop yield prediction using deep neural networks. *Frontiers in plant science*, 10:621, 2019.

[21] M. Khan, S. Ali, R. Ahmed, and A. Rashid. Smart fertilization systems using iot and ai. *Applied Sciences*, 14(3):3456, 2024.

[22] Nyayu Siti Khodijah, Muntoro Muntoro, Iwan Setiawan, Nabila Nabila, Irman Irman, and Vania Vania. Npk nutrient status in soil and leaves of piper nigrum l. under different growth criteria. *BIODIVERS-BIOTROP Science Magazine*, 3(1), 2024.

[23] A. Kumar, R. Singh, M. Gupta, and P. Sharma. Integrating uavs and machine learning for crop monitoring and disease detection. *Applied Sciences*, 13(5):2417, 2023.

[24] K. Lee, H. Park, and S. Kim. Quadratic models for describing soil ph and crop yield relationships. *Journal of Soil Science*, 30:78–90, 2024.

[25] S. Lee, J. Kim, H. Park, and D. Choi. Leveraging deep learning for crop disease diagnosis in smart agriculture. *Computers and Electronics in Agriculture*, 202:107089, 2024.

[26] H. Li, Z. Zhang, F. Wang, and X. Li. Federated learning-based crop yield prediction using multi-source data. *IEEE Transactions on Big Data*, 2023.

[27] Leslie Lipper and David Zilberman. A short history of the evolution of the climate smart agriculture approach and its links to climate change and sustainable agriculture debates. *Climate smart agriculture: Building resilience to climate change*, pages 13–30, 2018.

[28] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. PMLR, 2017.

[29] T. Nguyen, T. Pham, H. Tran, and D. Le. Hybrid machine learning models for enhancing crop yield prediction. *Journal of Smart Agriculture*, 15(3):76–88, 2024.

[30] R. Patel, S. Kumar, A. Singh, and A. Mishra. Ai-driven soil health monitoring using sensor networks. *Applied Sciences*, 13(4):1920, 2023.

[31] B Ramamoorthy and M Velayutham. The" law of optimum" and soil test based fertiliser use for targeted yield of crops and soil fertility management for sustainable agriculture. 2011.

[32] K. Roberts, J. Lee, and P. Wilson. Machine learning approaches for enhancing crop yield predictions. *Journal of Agricultural Technology*, 15(1):77–89, 2024.

[33] Aina Lundon Russenes, Audun Korsaeth, Lars R Bakken, and Peter Dörsch. Spatial variation in soil ph controls off-season n2o emission in an agricultural soil. *Soil Biology and Biochemistry*, 99:36–46, 2016.

[34] DRK Saikanth, Modi Ragini, Gagan Tripathi, K Rajesh, Anmol Giri, Shivam Kumar Pandey, and Lalit Kumar Verma. The impact of emerging technologies on sustainable agriculture and rural development. *International Journal of Environment and Climate Change*, 14(1):253–263, 2024.

[35] Uferah Shafi, Rafia Mumtaz, José García-Nieto, Syed Ali Hassan, Syed Ali Raza Zaidi, and Naveed Iqbal. Precision agriculture techniques and practices: From considerations to applications. *Sensors*, 19(17):3796, 2019.

[36] R. Singh, A. Arora, P. Gupta, and R. Kaur. Machine learning for predictive analysis in smart agriculture. *Journal of Smart Agriculture*, 12(2):98–110, 2023.

[37] J. Smith, R. Brown, and T. White. Iot-based real-time monitoring system for precision agriculture. *Journal of Agricultural Technology*, 18:123–134, 2022.

[38] J. Smith, K. Johnson, A. Lee, and B. Brown. Federated learning for climate-resilient agriculture. *Applied Sciences*, 14(4):4004, 2024.

[39] Behnam Tashayo, Afshin Honarbakhsh, Aliasghar Azma, and Mohammad Akbari. Combined fuzzy ahp–gis for agricultural land suitability modeling for a watershed in southern iran. *Environmental Management*, 66:364–376, 2020.

[40] P. Taylor, G. Harris, and S. Evans. Real-time crop monitoring using iot sensors and machine learning. *Journal of Precision Agriculture*, 25(4):345–358, 2023.

[41] S. Taylor, B. Smith, and R. Brown. Iot-based smart farming: Integration and interoperability challenges. *Computers and Electronics in Agriculture*, 2023.

[42] Thomas Van Klompenburg, Ayalew Kassahun, and Cagatay Catal. Crop yield prediction using machine learning: A systematic literature review. *Computers and electronics in agriculture*, 177:105709, 2020.

[43] J. Wang, X. Liu, Q. Li, and Y. Zhang. Smart irrigation system using iot and machine learning for precision agriculture. *Computers and Electronics in Agriculture*, 197:106896, 2023.

[44] R. Williams, D. Patel, and M. Garcia. Crop yield prediction using machine learning techniques. *Computers and Electronics in Agriculture*, 195:122–134, 2023.

[45] Naiqian Zhang, Maohua Wang, and Ning Wang. Precision agriculture—a worldwide overview. *Computers and electronics in agriculture*, 36(2-3):113–132, 2002.

[46] X. Zhang, Y. Liu, and Z. Chen. Regression models for predicting crop yield based on soil parameters. *Journal of Agricultural Data Science*, 25:123–136, 2023.

[47] X. Zhu, Y. Li, Q. Wang, and L. Zhao. Application of deep learning in precision agriculture: A review. *Applied Sciences*, 12(3):1623, 2022.