



FERTICAST – Data-Driven Fertilizer Optimization System Embedded with Rainfall Prediction

Hemant Kushwaha¹, Swechchha Maurya², Shreya Jaiswal³, Jyoti Gaur⁴

^{1,2,3}UG Scholar, Dept. of CSE, Babu Banarasi Das Institute of Technology & Management, Lucknow, India

⁴Assistant professor, Dept. of CSE Babu Banarasi Das Institute of Technology & Management, Lucknow, India

Emails: hemantkushofficial@gmail.com¹, swechchhamaurya492@gmail.com², jaisshreya30@gmail.com³, jyotigaur100@bbdnitm.ac.in⁴

Abstract

Fertilizer management is a major challenge in modern agriculture, especially due to changing weather conditions that directly affect soil nutrients and crop growth. Traditional fertilizer recommendations are usually fixed and do not consider factors like rainfall, which can wash away essential nutrients and reduce their effectiveness. This often leads to unnecessary fertilizer usage, increased costs for farmers, and environmental issues such as soil and water pollution. To address this problem, this paper introduces Ferticast, a smart and data-driven fertilizer recommendation system. The system uses a Random Forest Regression model to predict the required levels of nitrogen (N), phosphorus (P), and potassium (K) based on crop type and environmental conditions like temperature, humidity, and rainfall. In addition, a rainfall-aware module analyzes short-term weather forecasts to decide whether it is safe to apply fertilizers or if it should be delayed to avoid nutrient loss. The results show that the system provides reliable and adaptive recommendations, improving fertilizer efficiency compared to traditional methods. By combining machine learning with weather intelligence, Ferticast helps farmers make better decisions. Overall, the proposed system supports sustainable agriculture by reducing waste, improving crop productivity, and promoting smarter fertilizer usage.

Keywords: Machine Learning, Fertilizer Recommendation, Random Forest, Rainfall Prediction, NPK Estimation, Precision Agriculture, Sustainable Farming

1. Introduction

Agriculture continues to be one of the most important pillars of economic growth and food security, especially in countries like India where a large portion of the population depends on farming for livelihood. To maintain soil fertility and ensure high crop productivity, fertilizers play a crucial role by replenishing essential nutrients such as nitrogen (N), phosphorus (P), and potassium (K) (Hu et al., 2019; Yin et al., 2018). However, applying fertilizers is not as simple as following fixed guidelines—crop requirements vary based on environmental conditions, soil properties, and climate patterns. One of the major challenges faced in modern agriculture is the inefficient and untimely use of fertilizers. Farmers often rely on traditional practices or static recommendations that do not account for dynamic factors like rainfall, temperature, and humidity (Rao & Gupta, 2018). Excessive or poorly timed rainfall

can lead to nutrient leaching and runoff, causing fertilizers—especially nitrogen and potassium—to be washed away before plants can absorb them (Smith & Johnson, 2017; Patel & Mehta, 2019). Insufficient moisture can also reduce nutrient availability in the soil, resulting in reduced crop yield or increased fertilizer costs. This problem is important not only from an economic perspective but also from an environmental standpoint. Inefficient fertilizer use contributes to soil degradation, groundwater contamination, and ecological imbalance (Liu & Zhao, 2019). With increasing climate variability and unpredictable weather patterns, the need for a smarter and more adaptive fertilizer recommendation system has become more critical than ever (Wang & Li, 2021). To address these challenges, this research proposes Ferticast, a data-driven fertilizer optimization system that integrates machine learning



with rainfall prediction. It dynamically estimates NPK requirements using a Random Forest Regression model while analyzing short-term weather forecasts (Breiman, 2001; Kumar et al., 2020). By incorporating a rainfall-based decision mechanism, the system determines both the optimal quantity and the appropriate timing of fertilizer application, improving nutrient efficiency and supporting real-time decision-making. [1-3]

2. Literature Survey

Fertilizer recommendation has traditionally relied on soil testing and empirical knowledge to guide nutrient application. Studies highlight the importance of maintaining soil nutrient balance for improving crop yield (Hu et al., 2019; Kumar et al., 2016). However, these approaches are largely static and do not adapt well to changing environmental conditions like rainfall and temperature. With the rise of data-driven agriculture, machine learning techniques have been widely applied to predict crop yield and nutrient requirements. Supervised learning models effectively analyze soil and environmental factors (Kumar et al., 2020; Singh & Kumar, 2018). Among these, Random Forest is particularly effective due to its ability to handle complex and nonlinear agricultural data (Breiman, 2001). Ensemble methods further improve prediction accuracy (Srivastava & Jain, 2019; Chen & Wu, 2019), while deep learning models like LSTM capture temporal variations but require large datasets and high computational resources (Zhang & Li, 2020). Recent studies have introduced decision support systems and cloud-based platforms to improve accessibility for farmers (Patel & Joshi, 2020; Li et al., 2021). Rainfall is a critical factor affecting fertilizer efficiency due to nutrient leaching and runoff (Smith & Johnson, 2017; Patel & Mehta, 2019). However, most systems treat rainfall only as an input for prediction rather than as a decision-making constraint (Zhang et al., 2020; Wang & Li, 2021). Overall, current approaches focus on predicting nutrient values but do not ensure whether fertilizer application is suitable under given weather conditions. This creates a gap between prediction and real-world usability. To address this, the proposed Ferticast integrates machine learning with rainfall-aware decision logic, enabling accurate nutrient

estimation and timely fertilizer application.

3. Problem Statement

Traditional fertilizer recommendation methods rely on static nutrient charts that do not adapt to changing environmental conditions. These approaches fail to consider critical factors such as rainfall variability, which directly affects nutrient availability through leaching and runoff. As a result, farmers often apply fertilizers at inappropriate times, leading to inefficient nutrient usage, increased costs, and reduced crop productivity. Although existing machine learning-based systems can predict nutrient requirements, many of them do not integrate real-time rainfall analysis into the decision process. This limitation results in recommendations that are not fully reliable under dynamic weather conditions. Therefore, there is a need for a system that combines nutrient prediction with rainfall-aware decision making to ensure efficient, timely, and sustainable fertilizer application. [4-6]

4. Proposed System

Ferticast is a data-driven fertilizer recommendation system that helps farmers decide both the right amount and the right time for fertilizer application. It uses machine learning along with weather analysis to generate accurate nutrient recommendations based on crop and environmental conditions. The system takes inputs such as crop type, temperature, humidity, and rainfall, and predicts the required values of Nitrogen (N), Phosphorus (P), and Potassium (K) using a Random Forest model. Rainfall is incorporated into the system's decision process to improve recommendation reliability.

The working is based on two main steps:

- Predict NPK values using machine learning
- Check rainfall conditions before allowing fertilizer application

What makes Ferticast different is that it is not just a prediction system but also a decision-making system. Unlike traditional models that provide fixed fertilizer values, it adjusts recommendations based on real-time weather conditions and prevents fertilizer use during unsafe rainfall periods. Compared to the Eco-Fertilization model, Ferticast improves the approach by adding rainfall intelligence and dynamic decision support, making it more practical and suitable for



real-world agricultural use. Overall, the system provides a smarter way of fertilizer management by combining prediction with environmental awareness, helping reduce waste and improve efficiency.

5. Methodology

The proposed Ferticast system is derived from the Eco-Fertilization model and enhanced with rainfall-aware decision logic. It integrates machine learning-based nutrient prediction with environmental analysis to deliver adaptive fertilizer recommendations. Unlike traditional approaches that rely on static nutrient guidelines, the system dynamically adjusts recommendations based on climatic conditions, particularly rainfall, which directly influences nutrient availability and loss. The overall approach is structured into four key stages: dataset utilization, data preprocessing, model development using Random Forest Regression, and the operational workflow of the system. [7-10]

5.1.Dataset

The system utilizes the Crop Recommendation Dataset, a publicly available dataset obtained from Kaggle, which contains agricultural and environmental parameters relevant to nutrient prediction. The selected input features include crop type, temperature, humidity, and rainfall, while the output variables correspond to nitrogen (N), phosphorus (P), and potassium (K), as shown in Table 1.

Table 1 Dataset Features Used for NPK Prediction

Input Features	Output Features
Crop	Nitrogen (N)
Temperature	Phosphorus (P)
Humidity	Potassium (K)
Rainfall	

Rainfall is included as an input feature due to its influence on nutrient availability and environmental conditions, enabling the model to generate more realistic and adaptive fertilizer recommendations.

5.2.Data Preprocessing

The dataset undergoes preprocessing to ensure compatibility with the machine learning model and to

improve prediction accuracy.

- 1. Feature Reduction:** The original dataset contains multiple attributes, out of which seven relevant features are selected using feature selection techniques.
- 2. Data Cleaning:** Irrelevant entries and inconsistencies are removed to ensure data quality. [11-13]
- 3. Crop Encoding:** The categorical crop feature is converted into numerical form using encoding techniques.
- 4. Data Structuring:** The dataset is divided into independent variables (crop, temperature, humidity, rainfall) and dependent variables (Label_N, Label_P, Label_K).

5.3.ML Model: Random Forest Regression

Ferticast adopts Random Forest Regression for predicting soil nutrient values, as it effectively models complex and nonlinear relationships between environmental factors and crop nutrient requirements. The model is chosen due to its robustness, high prediction accuracy, and ability to reduce overfitting through ensemble learning. The system is configured with 50 decision trees ($n_{estimators} = 50$), and separate models are trained for each nutrient- Nitrogen (N), Phosphorus (P), and Potassium (K) to improve prediction precision. Each Random Forest model consists of multiple decision trees built on different subsets of the data and features. During prediction, each tree generates an independent output, and the final nutrient value is computed as the average of all tree predictions. This ensemble mechanism improves stability, reduces variance, and ensures reliable nutrient estimation under varying environmental conditions.

5.4.ML Model: Working Steps of the Proposed Method

The Ferticast system operates through a structured pipeline that integrates user input, environmental data, and machine learning-based prediction to generate nutrient recommendations.

- 1. User Input:** The user selects crop, state, and city.
- 2. Weather Data Extraction:** Environmental parameters such as temperature, humidity, and rainfall are retrieved using an API.

3. **Input Preparation:** The collected data is formatted and encoded to match the trained dataset structure.
4. **Model Execution:** Random Forest models are applied separately to predict: Nitrogen, Phosphorus, Potassium
5. **Prediction Aggregation:** Each model generates predictions using multiple trees, and the final output is computed as the average. [14-16]
6. **Output Generation:** The system provides predicted NPK values for the given crop and environmental conditions.
7. **Ferticast Enhancement:** The predictions are adjusted based on rainfall conditions to prevent nutrient loss and ensure effective fertilizer application

5.5. Proposed Algorithm: Ferticast

The proposed algorithm combines nutrient prediction with rainfall-based decision-making to ensure effective fertilizer application. It processes user inputs and environmental data to generate NPK recommendations, which are further validated based on rainfall conditions. This ensures that fertilizer is applied only under suitable conditions, improving efficiency and reducing nutrient loss. The step-by-step procedure of the proposed algorithm is presented in Table 2. [17-20]

6. System Architecture

Ferticast is designed as a modular and data-driven fertilizer recommendation system that integrates machine learning with real-time weather intelligence. The architecture follows a layered structure where user inputs are progressively transformed into adaptive fertilizer recommendations. The overall system architecture is illustrated in Fig.1, which presents the interaction between the major components of the system. The architecture is composed of four primary modules: Input Module, Weather Intelligence Module, Machine Learning Prediction Module, and Rainfall-Based Decision Engine. The process starts with user inputs, which are enhanced with environmental data and processed through a predictive model, then verified by a decision layer. This layered approach ensures

recommendations are both data-driven and environmentally practical. Figure 1.

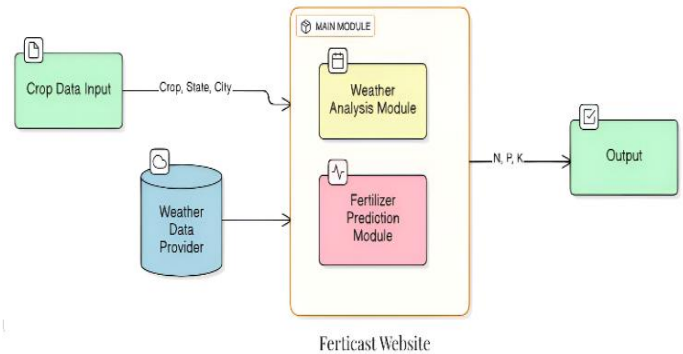


Figure 1 System Architecture of Ferticast

7. Module Description

Table 2 Step and Description

Step	Description
Step 1	Load the crop nutrient dataset and prepare it for model training.
Step 2	Apply preprocessing: encode crop names and normalize features if necessary.
Step 3	Select Random Forest Regression with n estimators = 50.
Step 4	Train three separate models: one each for Label_N, Label_P and Label K.
Step 5	Randomly generate multiple Decision Trees using different subsets of data and features.
Step 6	Each tree predicts nutrient values independently.
Step 7	Compute the final output as the average of predictions from all trees.
Step 8	Return the predicted N, P and K values for the selected crop and present weather conditions.

1. Input Module: The Input Module acts as the entry point of the system, collecting essential user inputs

such as crop type, state, and city. The module ensures that inputs are properly formatted and standardized before forwarding them to subsequent components. Its simple interface design enables ease of use for farmers and non-technical users. Table 2 shows Step and Description

2. Weather Intelligence Module: The Weather Intelligence Module integrates real-time climatic data into the system. It retrieves parameters such as rainfall, temperature, humidity, and precipitation probability using external APIs. The module focuses on short-term forecasts, typically spanning seven days, as these directly influence fertilizer application decisions. These parameters are processed and incorporated into the system to enhance the contextual relevance of fertilizer recommendations. By utilizing up-to-date weather information, the system ensures that recommendations are aligned with current environmental conditions and are more suitable for real-world agricultural scenarios.

3. Machine Learning Prediction Module: The Machine Learning Prediction Module is responsible for estimating nutrient requirements using a supervised learning approach. The system employs Random Forest Regression to model the relationship between crop type, environmental factors, and nutrient demand. The model takes crop, temperature, humidity, and rainfall as input features and predicts the required levels of nitrogen (N), phosphorus (P), and potassium (K). Separate models are trained for each nutrient to improve prediction accuracy. The ensemble nature of Random Forest allows the system to handle variability in agricultural data and reduces the risk of overfitting. This approach aligns with modern precision agriculture systems that use machine learning to capture complex interactions between soil, crop, and climate variables.

4. Rainfall-Based Decision Engine: The Rainfall-Based Decision Engine acts as the final validation layer of the system. It evaluates rainfall conditions before approving fertilizer application. Instead of directly displaying predicted values, the system first checks whether environmental conditions are suitable for fertilizer use. If rainfall exceeds predefined thresholds, the system either delays or blocks the recommendation to prevent nutrient loss due to

leaching and runoff. Under favorable conditions, the predicted NPK values are presented to the user. This module enhances the practical applicability of the system by converting it from a prediction tool into a decision support system, ensuring efficient fertilizer usage and minimizing environmental impact.

7.1. System Workflow and Data Flow

The Ferticast system processes data sequentially to generate fertilizer recommendations. Initially, the user provides crop and location details. These inputs are used to retrieve weather parameters such as temperature, humidity, and rainfall through an external API. The collected data is then passed to the machine learning model, which predicts the required NPK values based on crop and environmental conditions. The predicted values are further evaluated by the decision engine, which checks rainfall conditions before allowing fertilizer application. If rainfall is excessive or prolonged, the system delays or restricts the recommendation; otherwise, it displays the predicted nutrient values to the user. This ensures that recommendations are both data-driven and environmentally suitable. The flowchart of the system is illustrated in Figure 2.

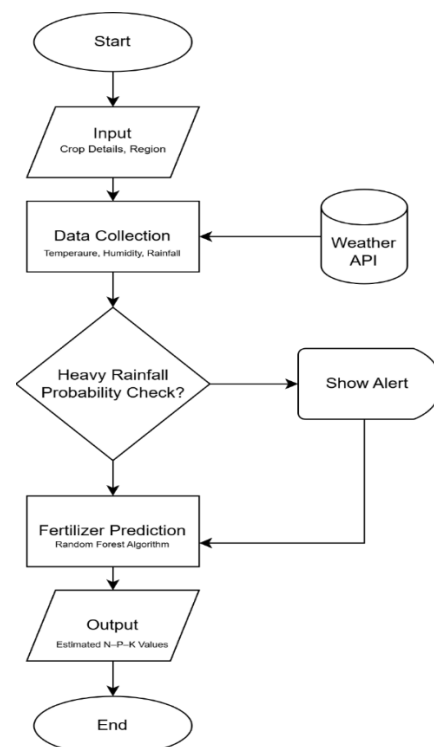


Figure 2 Data Flow Diagram Ferticast



8. Implementation

The implementation of the proposed system focuses on translating the conceptual design and methodology into a functional and deployable solution. This section outlines the development environment, dataset utilization, and data handling strategies adopted to ensure reliable model performance. Emphasis is placed on maintaining a modular and scalable structure while enabling seamless integration between data processing, machine learning prediction, and decision-making components [21]

8.1. Development Environment

The Ferticast system is implemented using Python due to its strong support for machine learning and data analysis. The Scikit-learn library is used for building the Random Forest Regression model, while Pandas and NumPy are used for data preprocessing and manipulation. The system is developed in Jupyter Notebook, enabling efficient experimentation and model evaluation. A modular approach is followed to separate data processing, prediction, and decision-making components, ensuring scalability and ease of maintenance.

8.2. Data Preparation and Train–Test Split

Prior to model training, the dataset undergoes preprocessing to ensure consistency and usability. Categorical variables such as crop type are encoded into numerical form, enabling their use in machine learning models. The dataset is then divided into training and testing subsets to evaluate model generalization:

- **Training set:** 80% of total data
- **Testing set:** 20% of total data

To align with the objective of climate-aware fertilizer recommendation, the dataset is refined by prioritizing weather-sensitive features. This ensures that the model captures the relationship between environmental conditions and nutrient demand more effectively. The split is performed using randomized sampling to avoid bias and ensure that both subsets represent the overall data distribution. This approach enables the model to learn patterns from the training data and validate its predictive capability on unseen samples. All preprocessing steps are applied consistently across both subsets to prevent data

leakage and maintain evaluation integrity.

8.3. System Integration

The implementation integrates data processing, prediction, and decision-making into a unified workflow. The system follows a sequential execution pipeline:

1. User inputs (crop and location) are collected through the interface
2. Weather parameters are retrieved dynamically via an external API
3. The processed inputs are passed to the trained model for prediction
4. Nutrient values (N, P, K) are generated as output
5. The decision engine evaluates environmental conditions before displaying results

This integrated approach ensures that predictions are not only data-driven but also context-aware, improving the reliability of fertilizer recommendations under varying climatic conditions.

9. Results and Discussion

The performance of the proposed system was evaluated using both train-test split (80:20) and k-fold cross-validation to ensure that the model generalizes well across unseen data. Cross-validation was particularly important in this work because agricultural datasets are highly variable due to changing weather conditions, and relying on a single split can lead to biased results. By validating the model across multiple folds, the system achieves more stable and reliable predictions. The Random Forest Regression model was assessed using standard regression metrics including R^2 (coefficient of determination), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). The model achieved an average R^2 score of around 0.85–0.88, indicating that it can effectively capture the relationship between crop type, environmental factors, and nutrient requirements. The MAE remained low across all nutrients, showing that the average prediction error is minimal, while RMSE values indicate that large deviations are limited and controlled. The model was trained separately for Nitrogen (N), Phosphorus (P), and Potassium (K), as each nutrient behaves differently under environmental conditions. Among these, Nitrogen



showed slightly higher variation due to its sensitivity to rainfall and leaching effects, whereas Phosphorus remained relatively stable, and Potassium exhibited moderate variability under high moisture conditions. Table 3 shows Fertilizer Recommendation Results

Table 3 Fertilizer Recommendation Results

Crop	Rainfall	N-P-K	Decision
Rice	Low	90-42-38	Apply
Wheat	Med	75-35-30	Delay
Cotton	High	82-40-36	Reject

The results demonstrate that the system not only predicts nutrient values but also applies rainfall-based decision logic to regulate fertilizer usage. Fertilizer application is permitted under low rainfall conditions, delayed under moderate conditions, and restricted during high rainfall periods to improve efficiency and reduce potential nutrient loss.

Conclusion

This research introduces Fercicast, a smart fertilizer advisory system that integrates machine learning prediction with climate-aware decision-making to address the limitations of traditional fertilizer recommendation methods. By leveraging Random Forest Regression, the system accurately estimates crop-specific NPK requirements, while the incorporation of rainfall intelligence ensures that these recommendations are applied under suitable environmental conditions. The proposed approach moves beyond static fertilizer charts by introducing a dynamic and adaptive framework, where nutrient recommendations are continuously influenced by real-time weather patterns. The evaluation results demonstrate high predictive performance and consistency, supported by cross-validation, confirming the model's ability to generalize across varying agricultural conditions. The rainfall-based filtering mechanism significantly enhances practical usability by preventing fertilizer application during periods of high leaching risk. The impact of Fercicast goes beyond accuracy by promoting sustainable agriculture through reduced fertilizer wastage and improved nutrient efficiency. By combining crop requirements with weather conditions, it enables

smarter and more responsible farming decisions. Its web-based design ensures easy access for farmers, making data-driven recommendations practical and scalable for real-world use.

Future Work

Fercicast can be further improved by integrating IoT sensors to collect real-time soil data such as moisture and nutrient levels, which would increase prediction accuracy. Developing a mobile application can enhance accessibility for farmers. The model can also be strengthened using larger and region-specific datasets and advanced techniques like deep learning. Additionally, incorporating satellite data and real-time alerts can make the system more accurate and practical for precision agriculture.

Acknowledgements

The author would like to sincerely thank the Babu Banarasi das Institute of Technology, Lucknow, and the Department of Computer Science for providing the academic support, technical resources, and guidance necessary to complete this project. The author is also grateful to the professors and colleagues whose insightful comments, encouragement, and constructive criticism greatly enhanced the quality and applicability of this research. Their collaboration was invaluable in successfully integrating machine learning techniques with practical applications. Special thanks are extended to Ms. Jyoti Gaur, Assistant Professor, Department of Computer Science, for her expert guidance, continuous support, and assistance throughout the course of this research project.

References

- [1] Hu, C., Zhang, Y., Li, Y., & Pan, G. (2019). Yield, nitrogen use efficiency and balance response to thirty-five years of fertilization in paddy rice–upland wheat cropping system. *Plant, Soil and Environment*, 65(2), 57–64.
- [2] Yin, Y., Ying, H., Zhang, H., & Li, X. (2018). Optimal NPK ratio estimation for rice cultivation under varying soil conditions. *Agricultural Sciences*, 9(12), 1012–1023.
- Smith, P., & Johnson, R. (2017). Impact of rainfall variability on soil nutrient leaching in cropping systems. *Journal of Soil and Water Conservation*, 72(3), 215–223.



- [3] Kumar, A., Singh, S., & Verma, P. (2016). Crop rotation and soil health: Long-term effects on nutrient dynamics. *Soil Biology & Biochemistry*, 95, 52–60.
- [4] Zhang, L., Chen, W., & Li, J. (2020). Machine learning applications for predicting crop yield and nutrient demand. *Computers and Electronics in Agriculture*, 175, 105589.
- [5] Patel, D., & Mehta, S. (2019). Rainfall-based predictive modeling for precision agriculture. *Agricultural Meteorology*, 267, 45–55.
- [6] Rao, M., & Gupta, R. (2018). Farmer decision-making and input use: Risk preference and fertilizer efficiency. *Agricultural Economics*, 49(5), 589–601.
- [7] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- [8] Li, Y., Zhang, X., & Wang, H. (2020). Deep learning for crop yield prediction under varying environmental conditions. *Computers and Electronics in Agriculture*, 172, 105335.
- [9] Srivastava, P., & Jain, A. (2019). Ensemble models for agricultural decision support systems. *Agricultural Systems*, 172, 19–30.
- [10] Chen, Y., Wang, X., & Liu, Z. (2021). Soil nutrient classification using automated machine learning techniques. *Geoderma*, 401, 115365.
- [11] Kumar, R., Singh, A., & Sharma, V. (2020). Integrating soil, crop, and weather data for smart fertilizer recommendations. *Precision Agriculture*, 21(6), 1234–1250.
- [12] Li, H., Zhao, Q., & Feng, Y. (2021). IoT-enabled fertilizer advisory systems: A review. *Computers and Electronics in Agriculture*, 187, 106290.
- [13] Zhang, W., & Li, M. (2020). LSTM networks for crop nutrient uptake modeling. *Neural Computing and Applications*, 32, 15719–15731.
- [14] Wang, J., Zhou, Y., & Hu, L. (2019). Remote sensing integration in precision nutrient management. *International Journal of Applied Earth Observation and Geoinformation*, 82, 101907.
- [15] Patel, K., & Joshi, R. (2020). Mobile and cloud-based decision support for fertilizer management. *Computers and Electronics in Agriculture*, 178, 105778.
- [16] Singh, N., & Kumar, P. (2018). Evaluating machine learning-based fertilizer recommendation models. *Agricultural Systems*, 162, 18–28.
- [17] Chen, L., & Wu, H. (2019). Performance comparison of ML algorithms in nutrient recommendation. *Precision Agriculture*, 20, 556–571.
- [18] Zhang, Y., Li, J., & Sun, Q. (2020). Integrating soil and rainfall data for resilient fertilizer advisory. *Agricultural Water Management*, 234, 106120.
- [19] Wang, X., & Li, Z. (2021). Seasonal forecasts and adaptive nutrient management systems. *Field Crops Research*, 261, 108018.
- [20] Liu, H., & Zhao, J. (2019). Optimizing fertilizer inputs for environmental and yield outcomes using ML. *Environmental Modelling & Software*, 117, 77–88.
- [21] Kaggle. Crop Recommendation Dataset. Retrieve from <https://www.kaggle.com/datasets/> (accessed 2025).