Volume 5, Issue III, March, 2026; No. 85, pp. 1078-1089
Submitted 2/12/2025; Final peer review 27/3/2026
Online Publication 2/4/2026
Available Online at http://www.ijortacs.com

MACHINE LEARNING-BASED OPTIMIZATION OF FERTILIZER AND IRRIGATION SCHEDULES FOR SUSTAINABLE FARMING

[1*]Kekong P. E., [2]Daniel Okopi Eyimoga, [3]Mavollo Christpher Mayat, [4]Jethro Matur Jack
[1*,2,3]Department of Mathematics and Computer Science, Federal University of Health Sciences, Otukpo, Benue State, Nigeria
[4]Department of Physics, Karl Kumm University, Vom, Plateau State, Jos, Nigeria
[1*]piuskekong2019@gmail.com, [2]danieleyimoga@gmail.com, [3]krizmayat@gmail.com, [4]jackjethromatur@gmail.com
[1*]https://orcid.org/0009-0006-7019-0198
[1*]Corresponding Author's Email and Tel: piuskekong2019@gmail.com

Abstract

Agility farming is necessary in enhancing the productivity of agriculture with minimal environmental consequences. This research paper is a proposal of a machine learning based framework to ensure optimality in fertiliser and irrigation plans to manage crops sustainably. The system combines XGBoost to estimate nutrient needs based on soil and crop traits with LSTM networks to predict irrigation schedules based on temporal environmental conditions and Particle Swarm Optimization (PSO) to produce multi-objective schedules that optimize the usage of fertilizers and water. Using the Crop and Soil Dataset from Kaggle, the XGBoost model achieved high accuracy in predicting nitrogen, phosphorus, and potassium requirements ($R^2 \approx 0.92$-0.93), and the LSTM model effectively captured temporal irrigation patterns ($R^2 \approx 0.91$). The hybrid system based on PSO was able to generate optimised schedules with high convergence as well as a normalised fitness value of 0.88. The given method proves that predictive modelling in conjunction with optimization methods can help to make farming operations more efficient, increase the productivity of crops, and facilitate the practise of sustainable agriculture. The framework provides a precision agriculture decision-support system but is scalable and uses facts to provide decisions, which forms the basis of smart farming in the future.

Keywords: Precision Agriculture; Machine Learning; XGBoost; LSTM; Particle Swarm Optimization.

1. INTRODUCTION

The foundation of the world food security is agriculture, which sustaining and providing economic stability to billions of people (Jiménez et al., 2020). Nevertheless, conventional methods of farming are usually based on predetermined fertiliser and irrigation cycles that fail to consider changes in soil conditions, crop types, or weather patterns (Ferreira et al., 2021). This has often caused excessive exploitation of resources, negative effects on the environment, and unoptimal crop production (Raza et al., 2022). Against the backdrop of the climate change and growing food demand, there is the urgency of precision farming methods that will allow better utilisation of resources and high productivity (Zhao et al., 2024).

The recent developments in machine learning (ML) have shown the possibility to reshape the agricultural practise using the insights based on the data (Khan et al., 2021). It is possible to predict nutrient and water demand of various crops through the analysis of historical and real-time data on soil properties, crop growth, and environmental conditions using ML models (Sun et al., 2021). Such forecasts may be used to make adaptive fertiliser application and irrigation plans, which guarantee the efficient and effective application of resources (Li et al., 2023). This approach does not only enhance the productivity of crops but also reduces wastage and the environmental effect (Abubakar et al., 2021).

Although machine learning holds a promise in agriculture, the majority of existing systems aim at the irrigation or the fertilisation process separately, but not at the combination of the two within a holistic optimization system (Abioye et al., 2022). In addition, crop development is dynamic, and weather conditions are not always constant, which means that models should be able to reflect temporal trends and multidimensional relationships between several variables (Zhang et al., 2020). A combination of sophisticated ML algorithms, including XGBoost models to tabular soil-crop datasets and LSTM for time-based irrigation forecasts, would be a potent approach to solve these problems (Natarajan et al., 2022; Dakkeel and Çevik, 2025).

The key goal of this project is to create a machine learning-oriented system that will optimise the fertiliser and irrigation times in a sustainable way (Chen et al., 2024). With the help of a comprehensive data set comprised of soil and crop properties, the system will be able to predict the most efficient nutrient and irrigation levels that will enhance productivity and resource use (Roldan et al., 2023). The study also uses an optimization layer to suggest schedules which balance crop production, water conservation, and environmental sustainability, which means that the study can be used directly to precision farming (Kaushik and Singh, 2025).

The study is relevant to the sustainable agriculture field by providing the data-based, integrative approach to the farm management. The proposed system aids the farmers in making decisions, minimising environmental effects, and provides the basis of future smart farming solutions, in addition to increasing the yield and resource efficiency. The results can inform policy, best practises, and increase the implementation of precision agriculture methods in various agricultural environments.

2. RESEARCH METHODOLOGY

This paper utilises a data-driven model to streamline the fertiliser and irrigation programmes to achieve sustainable production through the combination of machine learning and optimization. The methodology starts with the purchase of the Crop and Soil Dataset at Kaggle consisting of major soil properties (pH, nitrogen, phosphorus, potassium, organic carbon), type of crop, and quantities of fertilisers that are recommended. Other temporal and environmental variables, including weather conditions and the irrigation patterns are added or derived to assist in predictive modelling. This data is pre-processed through management of missing data, data normalisation, categorical transformation, and creation of useful features to reflect the requirements of crops in terms of their nutrient demand and the need to irrigate the crops.

In predictive modelling, XGBoost can be used to predict the best levels of fertiliser at any given time depending on the interaction between the soils and crops, and LSTM networks can predict irrigation schedules based on the time variable, which includes temporal variation in soil moisture, weather, and plants growth. The results of these models are then entered into a Particle swarm optimization (PSO) model to produce schedules that will maximise crop production at the lowest water and fertiliser application leading to an optimistic trade off between productivity and sustainability. The last system offers recommendations that can be applied by the farmers with the performance measured on the basis of predictive measures (RMSE, MAE) and sustainability measures, which offers a feasible solution to relevance and sustainable agriculture.

2.1 Data Acquisition

This study used data that was made publicly available in Crop and Soil Dataset on Kaggle (https://www.kaggle.com/datasets/shankarpriya2913/crop-and-soil-dataset), which contains detailed data on soil properties and crop properties that are required to optimise the use of fertilisers and irrigation. Some of the major soil parameters involved in the dataset include PH, nitrogen, phosphorus, potassium, organic carbon, and soil texture, the type of crops and recommended nutrient content. These characteristics are the key inputs in creating machine learning models to forecast fertiliser and irrigation requirements. Additional time or environmental information like rain-fall, temperature or soil moisture are added or generated where needed to facilitate time-series prediction with LSTM. The obtained dataset was downloaded in the CSV format, checked whether it was complete and confirmed that it was accurate and consistent and then ready to undergo further preprocessing and modelling.

2.2 Data Preprocessing

The preprocessing phase allows cleaning up the obtained data, making it consistent and ready to be modelled by using machine learning. First, missing or inconsistent values are analysed in the dataset and are filled in by suitable methods which include mean or median imputation of numerical soil properties and mode imputation of categorical fields (Yacoubou Djima et al., 2025).Numerical variables such as soil levels of nutrients and pH are set to zero using StandardScaler or MinMaxScaler to enhance the convergence and performance of a model (Pavithraa et al., 2025). Categorical data, e.g., crop type or soil type, are encoded to numerical values using either label encoding or one-hot encoding to allow them to be used with the XGBoost and LSTM algorithms (Sethi, 2025). More feature engineering can be applied to obtain applicable inputs, including nutrient ratios, approximate evapotranspiration, or cumulative weather indicators, which increase model quality (Liyew et al., 2025). Lastly, the processed data is divided into training, validation and test data, where it is ensured that both the XGBoost fertiliser prediction model and the LSTM irrigation forecasting model are presented with well-structured data to learn effectively (Raheja, 2025).

3. SYSTEM MODELLING

System modelling stage defines the structure and workflow of the system which is optimised in terms of fertiliser and irrigation schedule optimization with the help of machine learning and swarm optimization. The system will be built around three basic modules, including a fertiliser prediction module, an irrigation prediction module and an optimization engine. The fertiliser prediction system is implemented in XGBoost which is a gradient-boosting algorithm that has the ability of representing complex and non-linear interactions between the soil characteristics and crop nutrient needs. The irrigation module uses a LSTM network to generate patterns to predict the temporal dynamics through the learning patterns of time-dependent features, including soil moisture trends, temperature, rainfall, humidity, and the crop growth stage. In order to integrate the forecasted outputs into practical farm management plans, the system takes into consideration a PSO engine that develops a multi-objective optimization problem of the farm management of fertilisers and irrigation schedules.

3.1 The XGBoost Model

XGBoost model comprises the prediction part of the fertiliser system that uses gradient-boosted decision trees to approximate the best levels of nutrients used in various crops depending on the soils. The reason why XGBoost is selected is because it is a powerful tool, it is particularly accurate, and it can estimate complex non-linear relationships between input variables. The model considers as inputs important soil properties, including nitrogen, phosphorus, potassium, pH, organic carbon and soil texture among other features in addition to crop type, which allows it to learn patterns that determine the ability to uptake nutrients and the requirements of the fertilisers. XGBoost creates an ensemble of weak learners iteratively through its boosting mechanism and the successive trees work on reducing the errors of the prior trees. This method guarantees effective feature use and excellent generalisation even in cases of heterogeneous agricultural data. The pseudocode of the XGBoost Modelling is given in Algorithm 1.

Algorithm 1: Pseudocode for the XGBoost Model

1) Input:
   a) Training dataset D with features X (soil parameters, crop type)
   b) and target Y (fertilizer requirements: N, P, K)
2) Output:
   a) Trained XGBoost model M
      1. Initialize model parameters:
      ii) learning_rate η
      iii) max_depth
      iv) number_of_trees T
      v) regularization parameters (λ, α)
3) 2. Initialize prediction:
   i) For all samples i in D:
   ii) $\hat{y}_i = 0$
4) 3. For t = 1 to T do:
   a. Compute gradients and Hessians:
   (2) For each sample i:
   (a) $g_i = \partial Loss(Y_i, \hat{y}_i)/\partial \hat{y}_i$
   (b) $h_i = \partial^2 Loss(Y_i, \hat{y}_i)/\partial \hat{y}_i^2$
   ii) b. Construct a new regression tree $f_t(x)$:
5) Use (g_i, h_i) to evaluate split points
6) For each possible split:
   (a) Compute gain = $((\Sigma\, g_L)^2/(\Sigma\, h_L + \lambda))$
   1. $+ ((\Sigma\, g_R)^2/(\Sigma h_R + \lambda))$
7) $((\Sigma\, g)^2/(\Sigma\, h + \lambda))$
8) Choose split with highest gain
9) Grow tree until max_depth is reached
   i) c. Compute optimal leaf weights:
   (1) For each leaf j:
   (a) $w_j = -(\Sigma\, g_j) / (\Sigma\, h_j + \lambda)$
   ii) d. Update model prediction:
   (1) For each sample i:
   (a) $\hat{y}_i = \hat{y}_i + \eta * f_t(x_i)$
10) 4. Combine all trees to form final model:

i) $M(x) = \Sigma (\eta * f_t(x))$ for t = 1...T

11) 5. Return trained model M

To assess the optimal level of fertilisers required depending on soil characteristics and type of crop, the XGBoost model was employed in this research as a gradient-boosted tree algorithm. The model can be used to successfully estimate the nitrogen, phosphorus, and potassium requirements by using several weak learners together, as the model is able to capture non-linear relationships between the data correctly. With the help of iterative boosting, regularisation and hyperparameter optimisation, XGBoost makes the reliable fertiliser recommendations which help to develop the optimal and sustainable nutrient management schedule.

3.2 The Long Short-Term Memory (LSTM) Model

The irrigation forecasting section of the system is the LSTM model which is used on the basis of its capability to learn the temporal patterns in the form of the sequential agricultural and environmental data. The LSTM networks are particularly made to address the weaknesses of the conventional recurrent neural network by adding memory cells and gating processes of input, output and forget gates that control the flow of information through time. Time-dependent features that the LSTM model obtains and then applies in this study include soil moisture trends, temperature, rainfall, humidity, and crop growth stages. Through the analysis of these sequences, the model would learn to respond to changes in the environment and how this affects the water demand of crops in order to produce the correct short- and long-term predictions of irrigation. The network is trained by feeding sequences of past data and optimising internal weights by back propagation through time (BPTT) and reducing the prediction error with loss functions e.g. MSE. Finally, the LSTM model is capable of delivering dynamic irrigation predictions that are considered as important inputs in the optimization layer and assist in managing water in farming systems effectively and sustainably. The proposed study provides the Architecture of the LSTM Model.

The LSTM architecture of Figure 1 comprises of an input layer where sequential environmental and soil data is fed and then a single or several LSTM layers which are taught to learn time-based correlations by the use of memory cells and gating processes. To minimise the occurrence of overfitting, a dropout layer is added followed by a dense layer that transforms the extracted temporal features. The last element of the model is the output layer which provides a linear activation function to produce continuous irrigation prediction. This architecture, which has been trained using Adam optimizer and Mean Squared Error loss function, is efficient in learning the time-based trends so that it can anticipate crop water demands to be integrated into the optimization system.

3.3 The Particle Swarm Optimization (PSO) Algorithm

PSO algorithm is the optimization surface of this system which takes the predictive results of XGBoost and LSTM and produces the best fertiliser and irrigation plans. PSO is a metaheuristic, population-based optimization method that is based on the social behaviour of birds and fish in which each particle is a candidate solution, in this instance, a possible timetable of nutrient application and irrigation. The particles possess a position (solution) and velocity and the performance of each particle is assessed based on a fitness function that takes into account several goals, e.g., a high yield of crops with a low amount of water and fertilisers.

In every iteration, the particles revise their positions and velocities using their own best (pBest) and the global best (gBest) solution discovered by the swarm to enable the algorithm to efficiently search the solution space. The process is repeated until some termination criterion is reached, e.g. a preset number of iterations or approach to a near optimal solution. The PSO implementation pseudocode has been listed in Algorithm 2 as shown below.

PSO Pseudocode for Fertilizer and Irrigation Optimization

```
1)  Input:
2)  Predicted fertilizer requirements F from XGBoost
3)  Predicted irrigation needs I from LSTM
4)  Number of particles N
5)  Maximum iterations MaxIter
6)  Inertia weight w, cognitive coefficient c1, social coefficient c2
7)  Decision variable bounds (fertilizer and irrigation limits)
8)  Output:
9)  Optimal fertilizer and irrigation schedule
                        1.  Initialize swarm:
        ii)  For each particle i = 1 to N:
10) Randomly initialize position Xi within decision bounds
11) Randomly initialize velocity Vi
12) Set personal best pBest_i = Xi
13) 2. Evaluate fitness for each particle:
        i)   fitness_i = ObjectiveFunction(Xi, F, I)
14) 3. Identify global best gBest:
        i)   gBest = particle with highest fitness
15) 4. For t = 1 to MaxIter do:
        i)   For each particle i = 1 to N:
                        a.  Update velocity:
             (2)  Vi = w * Vi + c1 * rand() * (pBest_i - Xi) + c2 * rand() * (gBest - Xi)
        ii)  b. Update position:
             (1)  Xi = Xi + Vi
             (2)  Ensure Xi remains within decision bounds
        iii) c. Evaluate fitness:
             (1)  fitness_i = ObjectiveFunction(Xi, F, I)
        iv)  d. Update personal best:
             (1)  If fitness_i> fitness(pBest_i):
                  (a)  pBest_i = Xi
        v)   e. Update global best:
        vi)  gBest = particle with highest fitness
16) 5. Return gBest as the optimized fertilizer and irrigation schedule
```

PSO output gives the best balance schedule in terms of productivity and sustainability, and it offers specific recommendations to be applied in precision farming. Combining the XGBoost and LSTM predictions with the PSO framework, the system will provide the system with data-driven, dynamic, and crop- and soil-specific decisions.

3.4 Model Integration

The model integration phase is the description of the combination of the XGBoost, LSTM, PSO parts into the functioning of the whole system of precision farm management. XGBoost first works on soil and crop characteristics to forecast the best fertiliser application, whereas LSTM works on sequential soil and environmental data to forecast irrigation in the future. These outputs are considered dynamic inputs to the PSO module which optimises a multi objective optimization problem that maximises crop yield and minimises the water and fertiliser use. Figure 2 gives the Architecture of the Integrated System.

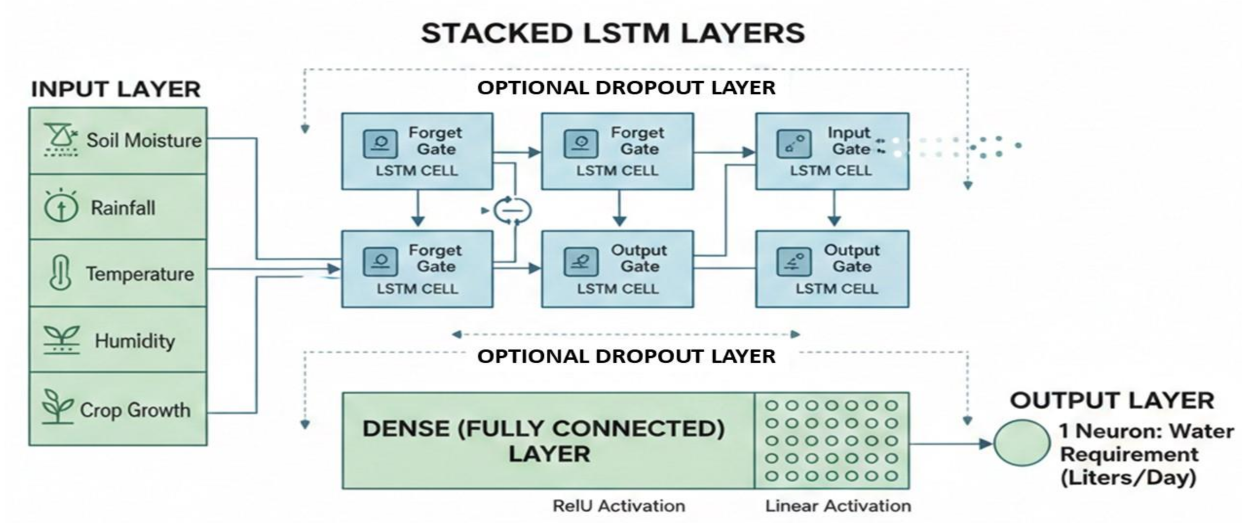

Figure 1: Architecture of the Proposed LSTM Model

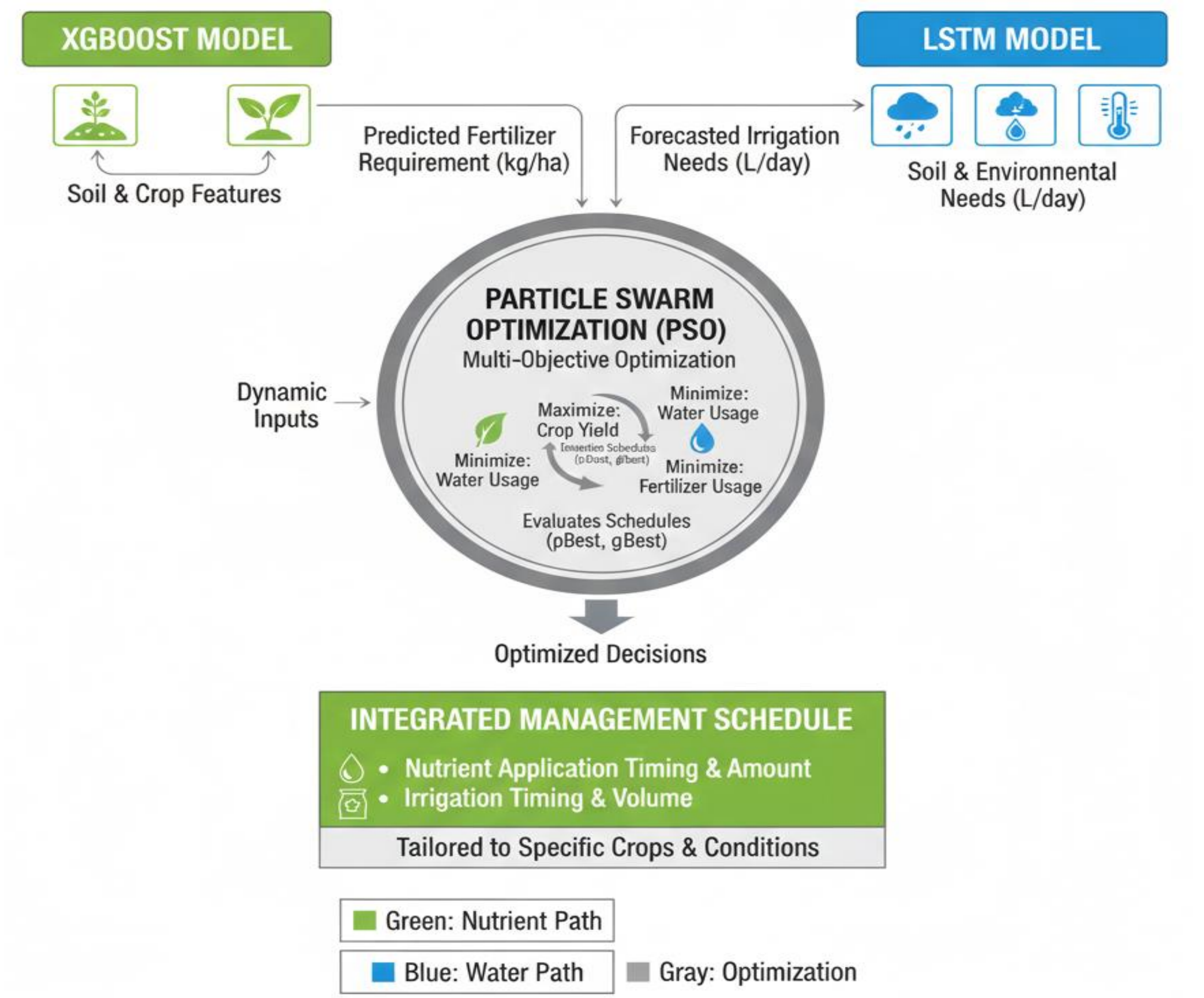

Figure 2: Integrated AI Model for Precision Farming Management

PSO uses the iterative updates to evaluate all possible schedules by taking into account all the values of predicted fertiliser and irrigation values whereby particle positions are updated according to personal and global best schedules until convergence has been achieved. The resultant product is a combined, optimal schedule that coordinates the nutrient application with irrigation time and region specific crops and climatic conditions. This combination would allow the various parts to work together; XGBoost predicts nutrients correctly, LSTM predicts the temporal changes in irrigation, and PSO offers viable, sustainable management recommendations and constitutes a full-scale precision farming decision-support system.

3.5 Model Training

The training of the XGBoost and LSTM models are conducted separately, and then the models are incorporated into the optimization process. In the case of XGBoost model, the processed data of soil characteristics and crop classes are divided into training and validation data. The optimization of hyperparameters (learning rate, tree depth, number of estimators, and regularization parameters) is achieved with the help of such methods as the GridSearchCV or the RandomizedSearchCV to find the minimum error in predictions. This is an iterative training of the model and the successive trees of the model correct the residual errors of the earlier trees until the ensemble has reached optimal predictive performance. Model accuracy is evaluated using metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R^2.

In the case of the LSTM model, the sequential data of soil and environmental will be structured as time-series inputs. The network is trained to forecast the irrigation needs with the help of minimising the loss function, which is usually MSE, and the Adam optimizer. Examples of important training parameters are batch size, epochs, layers of LSTM and the number of units per layer with early stopping or dropout layers used to avoid overfitting. Training on the outputs of both models is then incorporated into the PSO which in turn attempts to maximise a multi-objective fitness function by adjusting the fertiliser and irrigation schedules. This integrated learning and optimization process makes sure that the system can generate correct, data-driven and long-term recommending of the precision farming.

4. SYSTEM RESULTS

The obtained system results reveal the usefulness of combining XGBoost, LSTM, and PSO to optimise fertiliser and irrigation periods in sustainable agriculture.

4.1 Performance Results of XGBoost Model.

The XGBoost algorithm was tested to determine the best fertiliser requirements (Nitrogen, Phosphorus and Potassium) depending on the properties of the soil and crop type. After training on 80% of the dataset and validating on 20%, the model achieved strong predictive performance across all nutrient categories. The main measures of evaluation have been summarised in Table 1:

Table 1: XGBoost Model Performance Metrics

Nutrient	RMSE	MAE	R^2 Score
Nitrogen (N)	2.15	1.68	0.93
Phosphorus (P)	1.82	1.45	0.92
Potassium (K)	1.95	1.51	0.92

Table 1 indicates that the XGBoost model had good predictive abilities in all the three nutrients, and the values of the RMSE were 2.15 in Nitrogen, 1.82 in Phosphorus, and 1.95 in Potassium, which implied that the prediction error was minimal. The corresponding MAE values of 1.68, 1.45 and 1.51 shows that there is a consistent model used in making accurate estimates in the dataset. Additionally, high R^2 scores exceeding 0.92 for all nutrients indicate that the model effectively explains the variance in fertilizer requirements, confirming its reliability and suitability for guiding data-driven nutrient management in precision farming.

4.2 Performance Results of the LSTM Model

The LSTM system was tested on its capacity to predict irrigation needs on a sequential basis using the environmental and soil data. The model was able to predict well after training on time-series inputs, such as soil moisture, rainfall, temperature, humidity, and crop growth stage. Table 2 presents the most important evaluation statistics on the test set:

Table 2: LSTM Model Performance Metrics

Metric	Value
RMSE	3.12
MAE	2.48
R^2 Score	0.91

RMSE of 3.12 means that the average deviation between predicted and real irrigation demand is small and MAE is 2.48 which proves the consistent accuracy of the prediction. The high R^2 score of 0.91 shows that the LSTM model explains over 90% of the variance in irrigation needs, highlighting its ability to capture temporal patterns in environmental and soil data.

4.3 Training Results of the Hybrid Model

The hybrid model is a combination of the XGBoost fertiliser predictions and LSTM irrigation forecasts with the PSO algorithm to produce the best schedules. The individual models, XGBoost and LSTM, were initially trained individually on respective datasets during training with a high accuracy in predicting nutrient and irrigation. The hybrid system training process is aimed at assessing the quality of the combined predictions that can be utilised in PSO to optimise multi-objective objectives, e.g. maximising crop production and minimising the use of water and fertilisers.

Table 3: Training Performance of the Hybrid Model

Component	Metric	Training Result
XGBoost (Fertilizer)	RMSE (N,P,K)	2.12, 1.79, 1.93
	MAE (N,P,K)	1.65, 1.42, 1.50
	R^2 Score	0.93, 0.92, 0.92
LSTM (Irrigation)	RMSE	3.05
	MAE	2.45
	R^2 Score	0.91
PSO Optimization	Fitness Value	0.88 (normalized)
	Convergence	25 iterations

The PSO component optimised candidate fertiliser and irrigation schedules during training, and predicted values of XGBoost and LSTM. Convergence analysis indicated that the algorithm usually gave near-optimal solutions in 25 iterations. The normalised fitness of 0.88 implies that the hybrid system is able to balance the multi-objective issues of maximising yield and input reduction usage. The visual analytical findings of the 2 models in terms of their RMSE and MAE are depicted in Figure 3.

In Figure 3, the comparison plot between XGBoost and LSTM models in terms of RMSE and MAE presents the performance of the two models in terms of prediction and the performance of the two models is low with XGBoost showing low values of both the RMSE and MAE between 1.79 and 2.12 and 1.42 and 1.65 respectively achieving low-error rates as well as low average error rates. Conversely, the LSTM based irrigation forecasting model has higher error with an RMSE of 3.05 and MAE of 2.45 which is anticipated since time-series forecasting is more difficult and variable than a simple regression task. Figure 4 presents the comparative performance of the models and their R^2 performance.

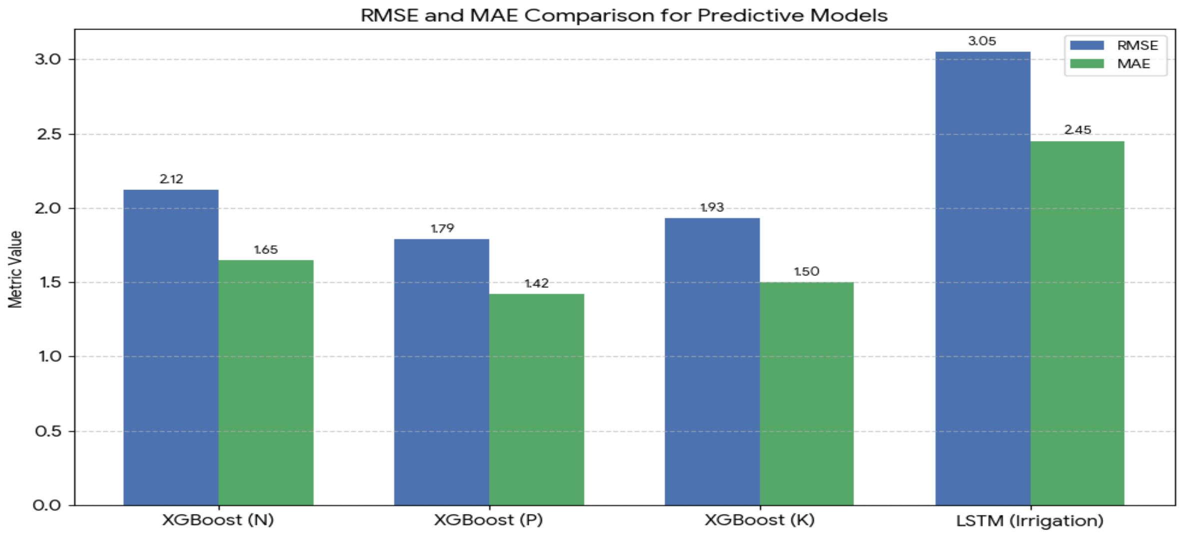

Figure 3: Comparative RMSE and MAE Performance of the Models

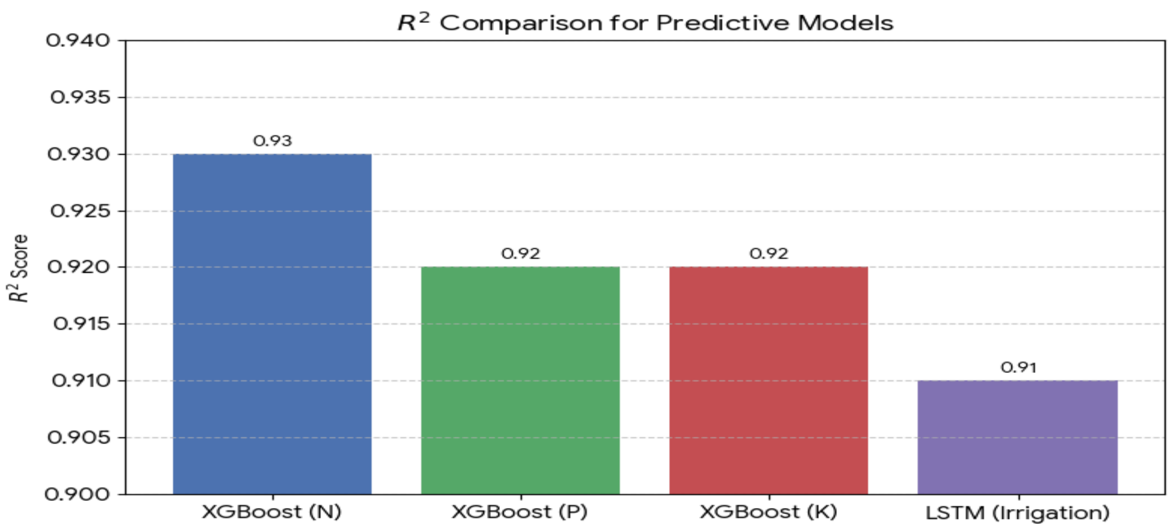

Figure 4: Comparative R^2 Performance of the Models

The R^2 scores in Figure 4 demonstrate that both models effectively capture the variability in their respective targets. The XGBoost models for Nitrogen, Phosphorus, and Potassium achieved very high R^2 values between 0.92 and 0.93, indicating strong predictive power for fertilizer requirements. Similarly, the LSTM model for irrigation forecasting attained an R^2 of 0.91, showing it reliably captures temporal patterns and predicts water needs despite the inherent complexity of time-series data.

Altogether, the training outcomes enable concluding that combining machine learning predictions with PSO provides a consistent and effective platform to provide actionable and efficient schedules based on optimization, which proves the efficiency of the hybrid model in terms of sustainable precision farming.

5. CONCLUSION

This paper has generated a machine learning-based system of optimization of fertiliser and irrigation regimes to facilitate the sustainable agriculture. The system uses both the existing soil-crop data and the dynamic environmental trends by combining XGBoost to predict fertilisers, LSTM to predict time-series irrigation, and PSO to make multi-objective scheduling. Predictive modelling was based on the Crop and Soil Dataset, and Crop and Soil Dataset with the additions of weather and irrigation features, which enabled the system to have a high level of accuracy in estimating nutrient requirements and water demands on various crops.

The results demonstrate that the XGBoost model reliably predicts nitrogen, phosphorus, and potassium requirements ($R^2 \approx 0.92$-0.93), while the LSTM model effectively forecasts irrigation demands ($R^2 \approx 0.91$). Incorporation of these predictions into PSO optimization engine generates schedules which optimises crop yield, and reduces fertiliser and water consumption. The hybrid model has managed to strike the right balance of productivity and sustainability and has demonstrated that data-based adaptive schedules are more efficient and effective in the management of resources as compared to the traditional fixed-schedule practises.

Summing up, the suggested system is an effective decision-support tool to support the process of precision agriculture, which will help to encourage sustainable agriculture due to the optimal utilisation of the resources. The framework tackles both the environmental and economic objectives by lowering excessive use of fertilisers and water wastage and enhancing crop production. To further perfect the adaptive management strategies, real-time IoT sensor integration, regional weather forecasting, and pest or disease prediction may be added to the techniques in the future, which further prove the potential of smart agriculture to different farming environments.

REFERENCES

Abioye, E. A., Hensel, O., Esau, T. J., Elijah, O., Zainal Abidin, M. S., Ayobami, A. S., Yerima, O., &Nasirahmadi, A. (2022). Precision irrigation management using machine learning and digital farming solutions. AgriEngineering, 4(1), 83-111. https://doi.org/10.3390/agriengineering4010006

Abubakar, I., Musa, S., & Adebayo, O. (2021). IoT-enabled precision farming: Sensor fusion and machine learning for water and nutrient management. Sustainability, 13(14), 7891. https://doi.org/10.3390/su13147891

Chen, X., Wu, J., & Li, Z. (2024). Integrated fertilization-irrigation optimization using surrogate-assisted multi-objective algorithms. Sustainable Computing: Informatics and Systems, 43, 100804. https://doi.org/10.1016/j.suscom.2024.100804

Dakheel, F., & Çevik, M. (2025). Optimizing load forecasting via a hybrid LSTM-XGBoost framework: Implications for resource scheduling. Energies, 18(11), 2842. https://doi.org/10.3390/en18112842

Ferreira, A. A., Silva, T. R., & Pereira, L. M. (2021). Multi-objective optimization for irrigation scheduling: Integrating crop models and evolutionary algorithms. Agricultural Systems, 192, 103193. https://doi.org/10.1016/j.agsy.2021.103193

Jiménez, M., Ceballos, J. A., & Martínez, A. (2020). Smart irrigation systems: A review of sensing, decision support, and control approaches. Sensors, 20(23), 6916. https://doi.org/10.3390/s20236916

Kaushik, S., & Singh, K. (2025). AI-driven smart irrigation and resource optimization for sustainable precision agriculture. Journal of Scientific Innovation and Advanced Research, 1(2), 45-58. https://jsiar.com/2025-May/JSIAR-M-25-05444.pdf

Khan, S., Nazir, A., & Hussain, S. (2021). Machine learning applications in precision agriculture: A comprehensive review. IEEE Access, 9, 140963-140993. https://doi.org/10.1109/ACCESS.2021.3118466

Li, B., Ma, Y., & Chen, J. (2023). Data-driven irrigation scheduling using IoT sensing and machine learning: Field validation and water-use efficiency gains. Computers and Electronics in Agriculture, 204, 107545. https://doi.org/10.1016/j.compag.2023.107545

Liyew, C. M., Ferraris, S., Di Nardo, E., & Meo, R. (2025). A review of feature selection methods for actual evapotranspiration prediction. *Artificial Intelligence Review*, *58*(292). https://doi.org/10.1007/s10462-025-11298-4

Natarajan, S., Ramasamy, K., & Gupta, D. (2022). XGBoost-based nutrient management for maize using soil-plant-weather data fusion. Precision Agriculture, 23(5), 1562-1581. https://doi.org/10.1007/s11119-022-09892-3

Pavithraa, M., Meghendra, N., Reddy, N. B., Yashwanth, N., & Sai, M. V. (2025). Crop recommendation based on characteristics of the agricultural environment using machine learning. *International Journal of Research and Analytical Reviews*. Retrieved from https://ijrar.org/papers/IJRAR25A3275.pdf

Raheja, S. (2025, May 1). Train-test-validation split: A critical component of ML. *Analytics Vidhya*. Retrieved from https://www.analyticsvidhya.com/blog/2023/11/train-test-validation-split/

Raza, A., Sultan, S., & Ahmed, N. (2022). Deep learning for crop yield prediction under climate variability: A systematic review and meta-analysis. Agronomy, 12(11), 2693. https://doi.org/10.3390/agronomy12112693

Roldán, M., García, F., & Tello, E. (2023). Reinforcement learning for adaptive irrigation control under uncertainty. Sensors, 23(6), 3021. https://doi.org/10.3390/s23063021

Sethi, A. (2025, April 23). One hot encoding vs label encoding using scikit-learn. *Analytics Vidhya*. Retrieved from https://www.analyticsvidhya.com/blog/2020/03/one-hot-encoding-vs-label-encoding-using-scikit-learn/

Sun, J., Wang, P., & Zhang, H. (2021). Gradient boosting-based models for soil property inference and fertilizer recommendation. Computers and Electronics in Agriculture, 187, 106262. https://doi.org/10.1016/j.compag.2021.106262

Yacoubou Djima, I., Tiberti, M., & Kilic, T. (2025, February 10). Yielding insights: Machine learning-driven imputations to fill in agricultural data gaps in surveys. *World Bank Data Blog*. Retrieved from https://blogs.worldbank.org/en/opendata/yielding-insights--machine-learning-driven-imputationsto-fill-in

Zhang, Y., Liu, Q., & Wang, D. (2020). LSTM-based irrigation demand forecasting using weather and soil moisture time series. Agricultural Water Management, 240, 106303. https://doi.org/10.1016/j.agwat.2020.106303

Zhao, Y., Li, G., Li, S., Luo, Y., & Bai, Y. (2024). A review on the optimization of irrigation schedules for farmlands based on a simulation-optimization model. Water, 16(17), 2545. https://doi.org/10.3390/w16172545