

Data Analytics Framework for Crop Yield Prediction Based on the Impact of Climate Using a Hybrid Model

Narinder Kaur^{1*}, Vishal Gupta²

NSUT East Campus (Formerly Ambedkar Institute of Advanced Communication Technologies and Research),
USICT, Guru Gobind Singh Indraprastha University, Delhi, India¹

NSUT East Campus (Formerly Ambedkar Institute of Advanced Communication Technologies and Research), New Delhi, India²

Abstract—In the modern Indian economy, agriculture is the most vital sector. Agriculture is an important source of revenue for the majority of the population of the world. Proper management of crops, crop species, crop field conditions, livestock, etc., can improve agricultural systems. The four primary subcategories of crop management are crop yield production management, weed management, crop disease management, and crop quality management. Planning agriculture, utilizing resources, and satisfying demand to ensure food security depend on accurate crop yield estimation. Crop yields are heavily influenced by climatic features like solar radiation, wind speed, rainfall, temperature, etc. Traditionally, crop yield management was completely performed by the farmers based on their experience using various techniques like crop rotation, mixed cropping, timely growing of crops, etc. These methods are combined with modern technologies like Artificial Intelligence, Internet of Things, drones, precision agriculture, etc., to improve crop productivity while considering preservation of natural resources. This study performed crop yield prediction using four baseline models based on machine learning: Decision Tree, Multiple Linear Regression, Support Vector Regression, and Random Forest. One of the best baseline models is further improved using nature-inspired algorithms. Four nature-inspired optimization algorithms used are: Whale Optimization Algorithm, Harris Hawks Optimization, Particle Swarm Optimization, and Grey Wolf Optimization. Particle Swarm Optimization-based optimized RF performed the best with $R^2=98.32$, $MAE=9.2766$.

Keywords—Crop yield prediction; impact of climate change; decision tree; support vector regression; linear regression and random forest; particle swarm optimization; whale optimization algorithm; Harris Hawks optimization; grey wolf optimization

I. INTRODUCTION

Climate change is an important challenge that impacts various living beings and our environment directly or indirectly. Continuous change in climate and the rise of CO₂ concentration in our environment are making agricultural system enhancement more challenging [1]. The Food and Agriculture Organization (FAO) reports that there has been a sharp rise in grain consumption and demand as compared to crop production. From 1964 to 2030, the demand for rice, wheat, and other coarse grains increased steadily [2].

In agriculture, food security is highly threatened by climate change, as this change affects crop production. The impact of climate change on agriculture includes sudden increases or decreases in temperature, variable precipitation patterns, and a

rise in the frequency of weather events [3][4]. In the agricultural sector, weather forecasting affects when crops are planted, harvested, and irrigated. This eventually helps to maximize crop yields and ensures food security [5]. Estimating crop production is a crucial component that depends on a number of variables, like soil characteristics, climate, and irrigation [6].

In the traditional crop assessment system, assessing soil health, crop growth, crop yield prediction, crop field management, etc. was done through manual observations, agricultural surveys, or by agricultural experts. However, wrong assessment led to reduced crop production. Numerous mitigating techniques have been employed over the past few decades, but the environment is still changing. Predicting crop yield is a challenge in modern precision agriculture, despite the proposal of numerous efficient models [7]. These models depict how temperature and precipitation variations directly impact crop production due to climate change.

The improvement of agricultural systems is greatly aided by Machine Learning (ML) approaches. ML is a scientific field that enables machines to learn without explicit programming [8]. Agriculture is increasingly using machine learning techniques to track and forecast environmental variables [9]. Numerous ML-based models have been proposed for agricultural system improvement [10]. However, managing massive volumes of data for precise conclusions and identifying the most significant climatic characteristics are the primary areas of worry. This can be resolved by employing an effective method that can also improve forecast performance.

In order to create more intelligent, quicker, and flexible systems, software engineers are increasingly using these natural designs. We refer to this method as "nature-inspired algorithms." It entails researching how organisms act, such as how ants locate food, how the brain learns, and how ecosystems maintain equilibrium, and then transforming those behaviours into functional software solutions. In order to solve open-ended problems, nature-inspired optimization strategies are both necessary and beneficial. Researchers and academicians can pursue a variety of avenues by comprehending nature-inspired methods and their significance for resolving practical issues.

A model for predicting agricultural crop yield while taking weather conditions into account has been proposed in this

*Corresponding author.

research. The following is the planned work's primary contribution:

- A complete data-analytic-based framework is created using existing algorithms for prediction of crop yield based on the impact of climate change through nature-inspired optimizing algorithms. The framework establishes how climate and agricultural variables can be incorporated into crop-yield prediction.
- Feature selection is performed using the RF approach.
- Crop yield prediction is performed using Multiple Linear Regression (MLR), Decision Tree (DT), Support Vector Regression (SVR), and Random Forest (RF) baseline models.
- The RF model is optimized using Particle Swarm Optimization (PSO), Whale Optimization Algorithm (WOA), Harris Hawks Optimization (HHO), and Grey Wolf Optimization (GWO) to predict crop yield.

This study's remaining sections are arranged as follows: Section II illustrates literature related to the proposed model. The suggested Crop Yield Prediction technique is briefly described in Section III. Section IV illustrates evaluation metrics used for comparison. Section V illustrates results and discussion. In Section VI, the entire study project is finally concluded.

II. RELATED WORKS

Numerous studies have been done in this field, which include identifying the best crop(s) and predicting crop yields using weather and soil information. Different articles related to the proposed work are summarized as follows:

Ramesh et al. [11] performed crop yield prediction using Neural Network (NN), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN), and Deep Neural Network (DNN) and proposed a hybrid model that combines DNN with WOA and Adaptive PSO, which performed better than the other models. Mohanty et al. [12] performed rice yield forecasting at the district level in Odisha using MLR and Quantile Regression (QR) as base models and proposed a hybrid model that merges RF, eXtreme Gradient Boosting (XGBoost), and Artificial Neural Networks (ANN) models based on crop and climate data of Odisha. Chaurasia et al. [13] proposed a hybrid model that combines LR, Support Vector Machines (SVM), k-Nearest Neighbors (kNNs), DT, RF, and XGBoost based on ensemble learning for crop yield forecasting.

Manjunath et al. [14] performed crop yield prediction using Random Forest (RF), XGBoost, Decision Tree (DT), SVM, and Linear Regression (LR) and then merged the three best of these models using the ensemble learning (Voting) technique and concluded that the hybrid model performed better. Dehghanisanij et al. [15] proposed a hybrid model that uses Adaptive Neuro-Fuzzy Inference System (ANFIS) and the Seasons Optimization (SO) algorithm to estimate water-use efficiency and yield in apple orchards. Cao et al. [16] performed wheat yield prediction at country level and field level using traditional models: RF and three deep learning-based models:

Deep Neural Networks (DNN), 1D Convolutional Neural Networks (1D-CNN), and Long Short-Term Memory (LSTM) model, and concluded that all four models performed well at country-level prediction, whereas DNN and RF performed too well at field-level prediction.

Khaki et al. [17] performed crop yield prediction of corn and soybean based on archived yield data, climate data, and soil data using a hybrid model that combines Convolutional Neural Network and Recurrent Neural Network (CNN-RNN). Han et al. [18] performed wheat yield prediction based on Climate, soil, vegetation index, cropland, and crop yield data using SVM, Gaussian process regression (GPR), RF, DT, NN, k-NN, boosted trees, and bagging trees, and concluded that SVM, RF, and GPR performed better than the other models. Sun et al. [19] proposed a hybrid model, CNN-LSTM, for prediction of Soybean yield based on past yield data, climate data, and Moderate Resolution Imaging Spectroradiometer (MODIS) data. Gopal et al. [20] proposed a hybrid model, MLR-ANN, to predict paddy yield based on Cropland, climate, soil, and past yield data. Hilal et al. [21] performed yield prediction of palm oil based on past yield data, cropland data, climatic data, and air pollutants data using Genetic algorithm/correlation analysis (GA/CA).

Sánchez et al. [22] performed crop yield prediction of multiple crops using Multiple linear regression, M5-Prime regression trees, perceptron multilayer neural networks, and concluded that M5-Prime and k-nearest neighbors performed better than other algorithms. To effectively predict agricultural productivity, it is crucial to observe climate patterns during sowing, the growing season, and throughout the year [23]. It is very difficult to accurately anticipate the crop growth stage using climate data because of very complicated dynamics that result in considerable changes in climatic factors [24]. Deep models used in [11], [17], [19], etc. have hidden parameter interactions. Stacking or voting architectures used in [12], [13], [14], etc. combine powerful baseline models but leave their hyperparameters fixed or unoptimized. In order to mitigate these concerns, this study has proposed an effective crop yield prediction model using RF optimization.

III. METHODOLOGY

Crop yield prediction while considering climate change is one of the important challenges in agriculture as climatic variables are fluctuating with respect to time. We depend on it to ensure food security and fulfill the future's food demand so as to accommodate the increase in the world's population of our world.

Fig. 1 depicts the methodology used for crop yield prediction. The proposed model is suitable for better prediction of yield, which can be used to increase crop yield production to fulfil future food demand. The model takes data related to crops and climate as an input. The input dataset is preprocessed by removing irrelevant columns and duplicate values. Then it is split into training data and Test Data. After splitting, One-Hot Encoding and normalization are performed in order to improve the classification accuracy. Following that RF based features selection is performed. The next step is to train the model using four machine learning baseline models: DT, MLR, SVR, RF and crop yield prediction is performed. Out of these four baseline

models, the best model is selected for further improvement using nature-inspired algorithms. Four RF optimized models: PSO optimized RF; WOA optimized RF; HHO optimized RF; and GWO optimized RF are trained and crop yield prediction is performed. The proposed model is explained as follows:

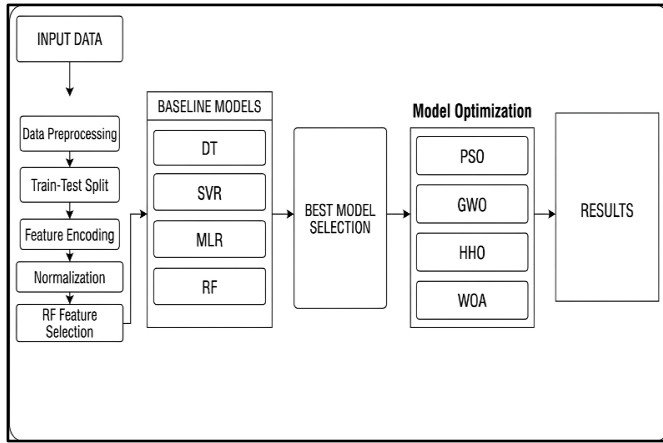


Fig. 1. The proposed model for crop yield prediction.

A. Crop and Climate Data

The crop yield dataset includes crop and climate data of different states of India taken from www.kaggle.com [36]. Dataset includes 19,689 unique samples/records. The dataset covers all states in India. The prediction based on the dataset can be used for improvement in the current agriculture system.

B. Data Preprocessing

In this part of the model, stateless preprocessing operations like removal of irrelevant columns and duplicate values, exploring dataset, separating the target variable from the others, etc., are done in this phase.

C. Train-Test Split

In the proposed model, two-way splitting is performed, splitting the data into training data and Test data. The model is fed the Train data in order to identify patterns and establish its own internal rules. The Test data serves as a last check to verify if the model can accurately predict data that it has never seen. In ML, to evaluate model performance, a Train-Test split is utilized. This split also avoids overfitting and helps in real-world simulation.

D. One-Hot Encoding and Normalization

A data preprocessing method called "One-Hot Encoding" is employed to convert categorical data into a numerical representation that is understandable by ML algorithms. For every category, it generates distinct binary columns, where 0 denotes the absence of a category, and 1 denotes its existence. The next step is to convert the encoded data into a form that is required by the data mining algorithms and improves the model. This conversion or transformation is known as data normalization. Here, Z-score normalization is used for data normalization [Eq. (1)].

$$Z = \frac{X - \mu}{\sigma} \quad (1)$$

where, X represents the actual value, μ represents the population's mean, and σ is the population's deviation. A correlation heat map is used for initial feature selection.

E. RF-Based Feature Selection

Feature selection is a process to select features that are most important for the model. By lowering overfitting, increasing accuracy, and cutting down on calculation time, it may improve the performance of the model. Feature selection, which involves figuring out those features which are most relevant for building a predictive model, is a crucial step in the machine learning process.

We conducted an ablation study by performing features selection using PSO, WOA, GWO, HHO, BACO (Binary Ant Colony Optimization) and RF. But PSO, WOA, HHO, GWO, and BACO selected features that excluded the main feature; no climate features were selected by these algorithms. Therefore, we used RF feature selection for the same, as it selected climate features, which justifies this manuscript.

F. Crop Yield Prediction

Crop Yield Prediction is performed using the features selected by RF through eight models: four baseline models DT, SVR, MLR, RF, and four optimized RF models, which are optimized using PSO, GWO, HHO, and WOA optimization algorithms. The following part of this section briefly explains these models. Training dataset with N samples and p features is expressible as:

$$Dataset = \{(X_i, Y_i)\}_{i=1}^N \quad (2)$$

where, $X_i = (X_{i1}, X_{i2}, \dots, X_{ik})$ i.e. i th sample of the *Dataset*, Y_i is the original yield of i th sample, N is the total count of samples, $k = 1, 2, \dots, p$ is a feature's index.

G. Decision Tree (DT)

A supervised learning-based approach used in ML, data mining, and statistics is decision tree learning. The DT approach can be used for classification as well as for regression. Here, it is used for regression purposes. DT approach creates a hierarchical tree structure that includes a single base node, branches, internal nodes, and leaf nodes. It functions similarly to a flowchart for step-by-step decision making [25] [26]. Fig. 2 depicts the Decision Tree Regression Approach briefly.

In DT, firstly, the mean of the node is computed as:

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i \quad (3)$$

where, $\bar{Y}$ represents the mean crop yield of the node, n represents total samples' count in the node, and Y_i is the actual crop yield of the i th sample in that node. After it Mean Squared Error is computed as:

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2 \quad (4)$$

The next step is to split, every feature is tested for different threshold values, and the weighted MSE is computed as:

$$MSE_{split} = \frac{|S_{Left}|}{|S|} MSE_{Left} + \frac{|S_{Right}|}{|S|} MSE_{Right} \quad (5)$$

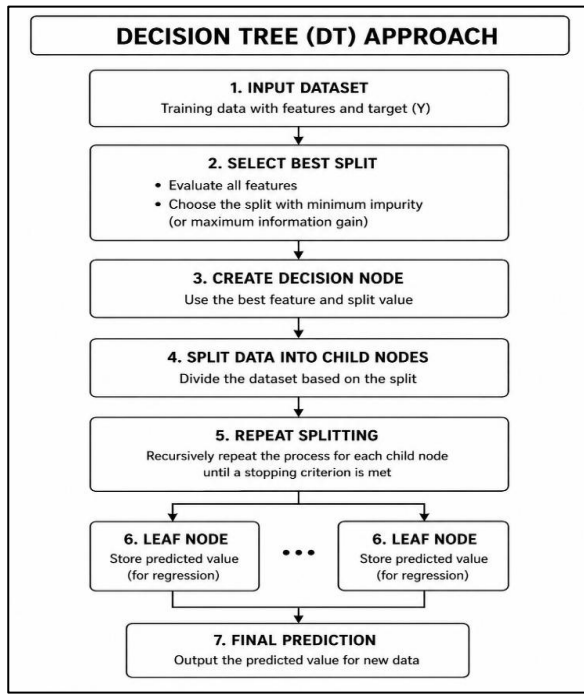


Fig. 2. Decision Tree (DT) working.

where, S is count of samples before splitting, S_{Left} is the count of samples in the left split, S_{Right} is count of samples in the right split, MSE_{Left} is the left split's Mean Square error, and MSE_{Right} is the right split's Mean Squared Error. MSE_{Left} and MSE_{Right} can be computed as:

$$MSE_{Left} = \frac{1}{S_{Left}} \sum (Y_{i_Left} - \bar{Y}_{Left}) \quad (6)$$

$$MSE_{Right} = \frac{1}{S_{Right}} \sum (Y_{i_Right} - \bar{Y}_{Right}) \quad (7)$$

where, Y_{i_Left} is the actual crop yield of a sample in the left node and Y_{i_Right} is the actual crop yield of a sample in the right node, $\bar{Y}_{Left}$ is mean crop yield of the left node, $\bar{Y}_{Right}$ is mean of crop yield of the right node. The best split is selected as:

$$(k^*, t^*) = \arg \arg MSE_{split}(k, t) \quad (8)$$

where, k is candidate's feature index, t is the candidate's threshold value, k^* is the best feature which is selected for split, t^* is the best feature's threshold, $MSE_{split}(k, t)$ is weighted MSE computed using j feature and t threshold value. Splitting is done till we reach a leaf node. Once a new sample reaches a leaf node, compute prediction as:

$$\hat{Y} = \frac{1}{n} \sum_{i=1}^n Y_i \quad (9)$$

where, $\hat{Y}$ is the predicted crop yield of the node, n represents total samples' count in the leaf node, Y_i actual crop yield of each sample in the leaf.

The reduction of the Mean Squared Error (MSE) is the primary goal of DT:

$$MSE = \frac{1}{N} \sum_{i=1}^N (Y_i - \hat{Y}_i)^2 \quad (10)$$

where, N is the training sample count, Y_i is actual crop yield of i th sample, $\hat{Y}_i$ is the predicted crop yield of the i th sample, $(Y_i - \hat{Y}_i)^2$ is the residual error squared.

H. Multiple Linear Regression (MLR)

MLR Targets Defining a Linear Relationship between a Dependent Variable and Numerous Independent Variables. Researchers Can Investigate the Effect of Multiple Independent Variables on the response simultaneously by Using Multiple Regression [27]. Fig. 3 depicts the working of the MLR Model.

MLR equation for crop yield of a sample is as follows:

$$Y_i = f(X_{i1}, X_{i2}, \dots, X_{ik}) \quad (11)$$

Eq. (11) can be converted to the following form:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \epsilon_i \quad (12)$$

where, β_0 is the intercept, $\beta_1, \beta_2, \dots, \beta_k$ are the coefficients of features ($X_1, X_2, \dots, X_k$) and ϵ_i represents the random error term of i th sample $i = 1, 2, \dots, N$.

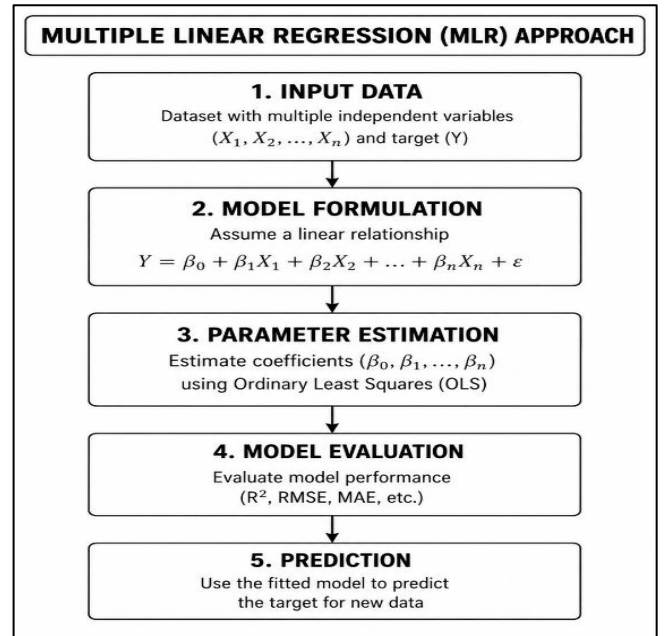


Fig. 3. Multiple Linear Regression (MLR) working.

The complete matrix representation of the above Eq. (12) can be written as:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & \dots & X_{1k} \\ 1 & X_{21} & \dots & X_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \dots & X_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix} \quad (13)$$

Ordinary Least Squares (OLS) solution is used to compute the regression coefficients as:

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (14)$$

MLR equation for crop yield prediction of a sample is as follows:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + \dots + \hat{\beta}_n X_{ik} \quad (15)$$

Residual error (e) is computed as:

$$e = Y_i - \hat{Y}_i \quad (16)$$

The reduction of the Mean Squared Error (MSE) is the primary goal of MLR also [same as of DT, Eq. (10)].

I. Support Vector Regression (SVR)

In 1996, Vladimir N. Vapnik, Harris Drucker, Christopher J. C. Burges, Linda Kaufman, and Alexander J. Smola introduced a regression using SVM. We refer to this technique as support vector regression (SVR). SVR aims at finding the optimal fit that precisely predicts the objective variable while lowering complexity to prevent overfitting. This is accomplished by defining an ϵ -insensitive tube around the hyperplane that allows for a certain amount of variance between the actual and anticipated target values. A machine learning method called non-linear Support Vector Regression (SVR) is used to forecast continuous values in situations where data associations are complicated or curved [28]. Fig. 4 represents the working of the SVR approach.

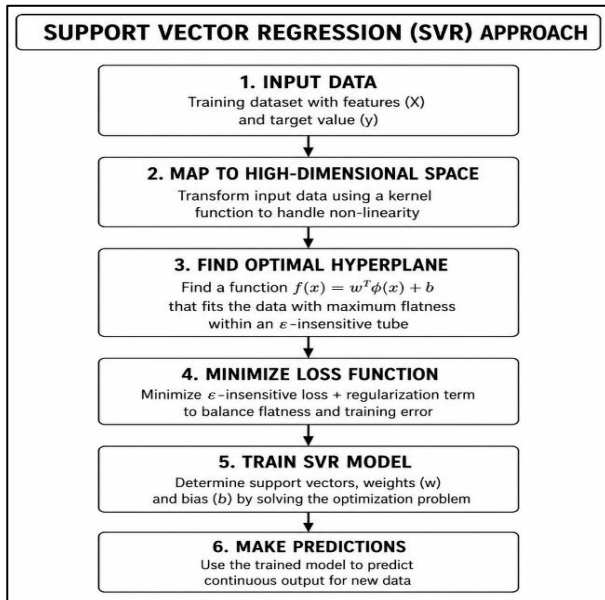


Fig. 4. Support Vector Regression (SVR) working.

The SVR non-linear function for prediction:

$$f(X) = \hat{Y} = W^T \phi(X) + \gamma \quad (17)$$

Where W^T represents the transpose of weight vector W , $\phi(X)$ is nonlinear mapping to a higher-dimensional feature space, γ is the bias.

SVR ignores small errors, known as error-insensitive, which can be written as:

$$L_\epsilon(Y, f(X)) = \begin{cases} 0 & |Y - f(X)| \leq \epsilon \\ |Y - f(X)| - \epsilon & |Y - f(X)| > \epsilon \end{cases} \quad (18)$$

Where, ϵ is the width of the error-insensitive tube. SVR objective is to optimize (minimize):

$$\min \frac{1}{2} \|W\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \quad (19)$$

with the following conditions:

$$Y - W^T \phi(X_i) - b \leq \epsilon + \xi_i \quad (20)$$

$$W^T \phi(X_i) + b - Y_i \leq \epsilon + \xi_i^* \quad (21)$$

$$\xi_i, \xi_i^* \geq 0 \quad (22)$$

Where, $\|W\|^2$ is Euclidean norm of W , C is the regularization parameter, ξ_i is error above the boundary, and ξ_i^* is error below the boundary.

Eq. (19) can be converted to a dual optimization problem using Lagrange multipliers as:

$$\max -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(X_i, X_j) - \epsilon \sum_{i=1}^N (\alpha_i + \alpha_i^*) + \sum_{i=1}^N Y_i (\alpha_j - \alpha_j^*) \quad (23)$$

with conditions:

$$0 \leq \alpha_i, \alpha_i^* \leq C \quad (24)$$

$$\sum_{i=1}^N (\alpha_i - \alpha_i^*) = 0 \quad (25)$$

where, i is first sample index, j is the second sample index, α_i is Lagrange multiplier which is associated with upper error-insensitive constraint for i th sample, α_i^* is Lagrange multiplier which is associated with the lower error-insensitive constraint for i th sample, $K(X_i, X_j)$ is the function of kernel that measures the similarity between the i th and j th samples in the transformed feature space, X_i is i th training sample, X_j is j th training sample, ϵ defines the width of the insensitive tube, Y_i is the target value, C is the regularization parameter.

Kernel function used in equation is expressed as:

$$K(X_i, X_j) = \exp \left(-\gamma \sum_{k=1}^p (x_{ik} - x_{jk})^2 \right) \quad (26)$$

The final prediction equation of SVR is as follows:

$$f(X) = \sum_{i=1}^N (\alpha_i - \alpha_i^*) \exp(-\gamma \|X_i - X\|^2) + b \quad (27)$$

J. Random Forest (RF) Regression

RF approach is an ensemble learning approach that creates numerous decision trees and uses bagging to mix their outputs to prevent overfitting and improve prediction accuracy. Bagging and boosting ensemble learning are the two primary categories of ensemble learning techniques. In bagging ensemble learning, distinct sub-datasets are generated from the primary dataset and sent to various machine learning models. The final output or outputs are created by combining the findings of these simultaneous models. In boosting, several machine learning models use the same dataset. These models operate in a sequential manner. The first model's findings are improved by the second model, the second model's findings are improved by the third model, and so on. Bagging is the foundation of the RF method. The Random Forest automatic machine learning technique was developed in 2001 by Leo Breiman and Adèle Cutler. [29]. Fig. 5. depicts working of the RF model.

In RF regression, Multiple trees T are used in RF approach, and for each tree, a separate Dataset D_t is created. Creating of different datasets is called bootstrap sampling and can be written as:

$$D_t \subseteq D, \text{ for } t = 1, 2, \dots, T \quad (28)$$

where, D_t is bootstrap dataset for the t^{th} tree, and T stands for total count of trees.

Each tree predicts the output i.e., crop yield independently as:

$$\hat{Y}_t = f_t(X) \quad (29)$$

where, $\hat{Y}_t$ is the prediction of t^{th} tree, X is input feature vector.

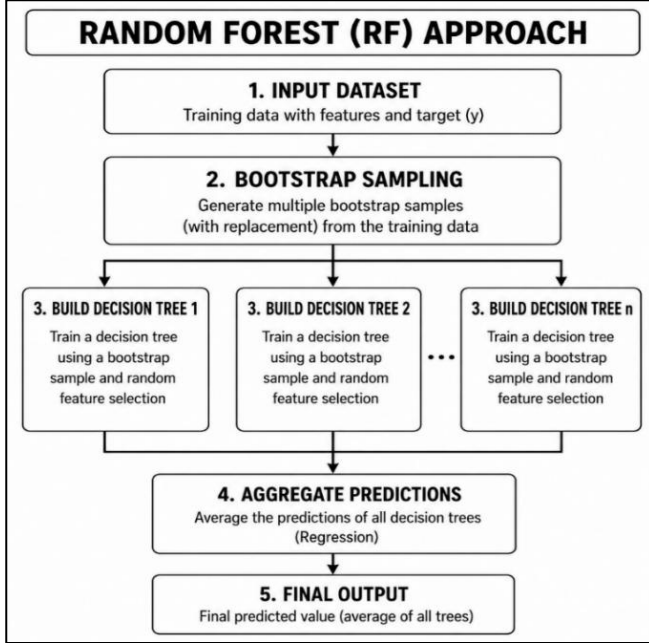


Fig. 5. Random Forest (RF) working.

The final RF Prediction equation can be written as:

$$\hat{Y} = \frac{1}{T} \sum_{t=1}^T f_t(X) \quad (30)$$

where, $\hat{Y}$ final predicted crop yield, T stands for the count of trees, $f_t(X)$ is predicted value of the t^{th} tree.

K. Optimized RF Models

Here, four nature-inspired optimization approaches are used to optimize the RF model: PSO, GWO, HHO, and WOA algorithms. These optimization algorithms find the best parameters for the RF model.

Let d be the number of parameters of RF model to be optimized; then initialize the Lower Bound (LowB) and Upper Bound (UpB) vectors to specify the lower and upper limits of each RF parameter that is to be optimized as:

$$\text{LowB} = [\text{LowB}_1, \text{LowB}_2, \dots, \text{LowB}_d] \quad (31)$$

$$\text{UpB} = [\text{UpB}_1, \text{UpB}_2, \dots, \text{UpB}_d] \quad (32)$$

Initialization of position matrix X is as follows:

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} X_{11} & \dots & X_{1d} \\ X & \dots & X_{2d} \\ \vdots & \ddots & \vdots \\ X_{n1} & \dots & X_{nd} \end{bmatrix} \quad (33)$$

where, count of rows = n represents no. of swarms (in PSO)/ wolves (in GWO)/ hawks (in HHO)/ whales (in WOA), count of columns = d represents parameters of the RF model to be optimized. A single row represents a single particle (in PSO)/ wolf (in GWO)/ hawk (in HHO)/ whale (in WOA) as $X_i = [X_{i1}, X_{i2}, \dots, X_{id}]$. Initialization of every value in the above matrix is done as follows:

$$X_{ij} = \text{LowB}_j + r_{ij}(\text{UpB}_j - \text{LowB}_j) \quad (34)$$

where, X_{ij} is the single value of the matrix shown in equation, UpB_j is the Upper Bound value of parameter j , LowB_j is the Lower Bound value of parameter j , r_{ij} is a random variable such that $0 \leq r_{ij} \leq 1$.

Eq. (31), (32), (33), and (34) are common in all four algorithms and belongs to the initialization part of it.

These four optimized RF models compute the fitness of a single row as:

$$\text{Fitness}(X_i) = \text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y}_i)^2} \quad (35)$$

L. PSO Optimized RF

PSO was first identified in 1995 by Kennedy and Eberhart. PSO technique is used to optimize continuous nonlinear functions. PSO is a stochastic population-based optimization method inspired by natural swarm intelligence. Every particle travels across the search space and represents a possible solution. Both individual experience and collective swarm knowledge direct movement. The approach uses a fitness function to iteratively improve solutions [30][31]. Fig. 6 represents the working of PSO-based RF optimization.

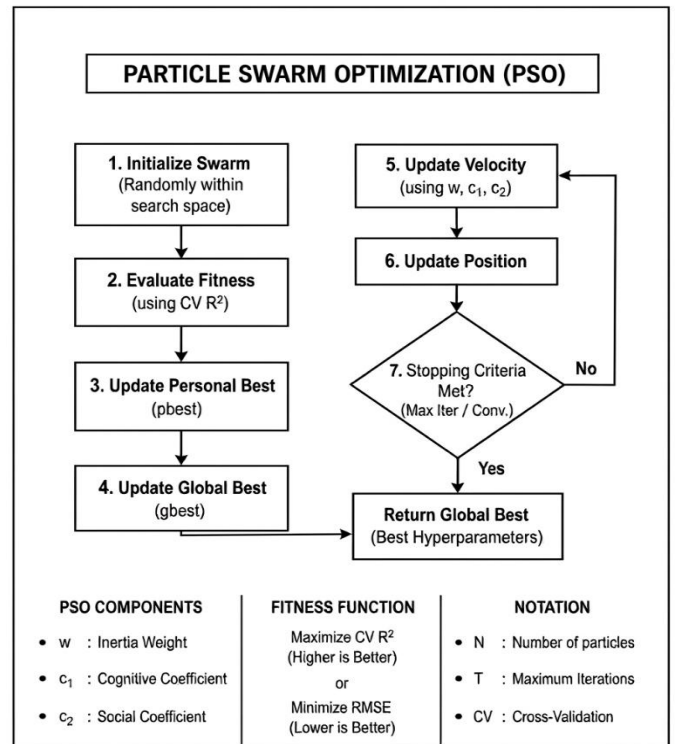


Fig. 6. Particle Swarm Optimization (PSO) working.

Steps to optimize RF model using PSO:

Initialization:

Initialize Lower Bound (LowB) and Upper Bound (UpB) vector: using Eq. (31) and (32)

Initialization of the swarm position matrix: using Eq. (33) and assign values to it using Eq. (34)

Initialize the velocity matrix

$$V = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix} = \begin{bmatrix} V_{11} \cdots V_{1d} \\ V_{21} \cdots V_{2d} \\ \vdots \quad \ddots \quad \vdots \\ V_{n1} \cdots V_{nd} \end{bmatrix} \quad (36)$$

Compute Fitness of every particle using Eq. (35)

Personal best is updated as:

$$PB_i = \begin{cases} X_i, & \text{if } \text{Fitness}(X_i) > \text{Fitness}(PB_i) \\ PB_i, & \text{otherwise} \end{cases} \quad (37)$$

where, PB_i best solution found by particle i

Global best's update as:

$$GB = \arg \max_i \text{Fitness}(P_i) \quad (38)$$

where, GB is the best solution found by entire swarm.

Update Velocity and particle position:

Velocity update:

$$V_i^{t+1} = \omega V_i^t + c_1 r_1 (PB_i - X_i^t) + c_2 r_2 (GB - X_i^t) \quad (39)$$

where, ω represents inertia weight, c_1 is cognitive coefficient, c_2 is the social coefficient, t represents current iteration.

Particle position update:

$$X_i^{t+1} = X_i^t + V_i^{t+1} \quad (40)$$

Repeat steps 2 to 5 until convergence is attained or the maximum number of iterations is attained.

RF is optimized with the best solution as,

$$X_{best} = GB \quad (41)$$

M. GWO Optimized RF

A nature-inspired metaheuristic algorithm called Grey Wolf Optimization (GWO) that imitates the leadership organization and hunting techniques of grey wolves in the wild. Seyedali Mirjalili first presented it in 2014. The best fit, according to GWO, is the Alpha wolf (α); the second best fit is the Beta wolf (β), the third best fit is the Delta wolf (δ); and the Omega wolves (ω) are the remaining possible options [32][33]. Fig. 7 depicts the working of GWO Based RF optimization.

Steps to optimize the RF model using GWO:

Initialization:

Initialize Lower Bound (LowB) and Upper Bound (UpB) vector: using Eq. (31) and (32)

Initialization of the wolf position matrix: using Eq. (33) and assign values to it using Eq. (34)

Compute Fitness for every wolf using Eq. (35)

Identify the Alpha (α), the Beta (β), the Delta (δ) and Omega (ω) wolves as:

$$X_\alpha = X_i \text{ with best Fitness}(X_i) \text{ value} \quad (42)$$

$$X_\beta = X_i \text{ with second best Fitness}(X_i) \text{ value} \quad (43)$$

$$X_\delta = X_i \text{ with third best Fitness}(X_i) \text{ value} \quad (44)$$

$$X_\omega = X_i \text{ with rest of the wolves} \quad (45)$$

Compute parameter a as:

$$a = 2 - \frac{2t}{T} \quad (46)$$

where, t shows the current iteration, and T stands for maximum iteration. Now compute the coefficient as:

$$A = 2ar_1 - a \quad (47)$$

$$C = 2r_2 \quad (48)$$

where, A and C are coefficients, where r_1 and r_2 are random values such that $0 \leq r_1, r_2 \leq 1$

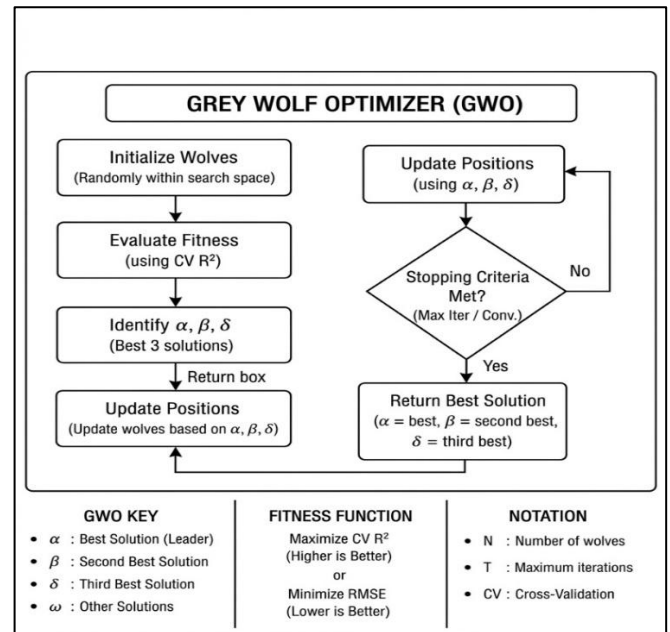


Fig. 7. Grey Wolf Optimization (GWO) working.

Compute the distance from X_α , X_β , and X_δ wolves and Compute candidate positions based on D_α , D_β , and D_δ as:

Distance from X_α , X_β , and X_δ wolves as:

$$Distance_\alpha = |C_1 X_\alpha - X| \quad (49)$$

$$Distance_\beta = |C_1 X_\beta - X| \quad (50)$$

$$Distance_\delta = |C_1 X_\delta - X| \quad (51)$$

Candidate positions based on D_α , D_β , and D_δ as:

$$X_1 = X_\alpha - A_1 \text{Distance}_\alpha \quad (52)$$

$$X_2 = X_\beta - A_2 \text{Distance}_\beta \quad (53)$$

$$X_3 = X_\delta - A_3 \text{Distance}_\delta \quad (54)$$

Update Wolf Position as:

$$X_{t+1} = \frac{X_1 + X_2 + X_3}{3} \quad (55)$$

Eq. (55) will update every wolf except Alpha wolf.

Check the updated position for its upper bounds and lower bounds as

$$\min(\text{UpB}, \max(\text{LowB}, X)) \quad (56)$$

Evaluate Fitness again and Update X_α , X_β , and X_δ again

Until the maximum count of iterations is attained, repeat steps 2 through 7.

RF is optimized with best solution as:

$$X_{\text{best}} = X_\alpha \quad (57)$$

N. HHO Optimized RF

Ali Asghar Heidari and associates introduced the well-known population-based, nature-inspired HHO metaheuristic algorithm in 2019. It imitates Harris's hawks' coordinated foraging and "surprise pounce" hunting techniques. Exploration, escaping energy shifts, and exploitation tactics are important elements of the HHO optimizer [34]. Fig. 8 represents the working of HHO-based RF optimization.

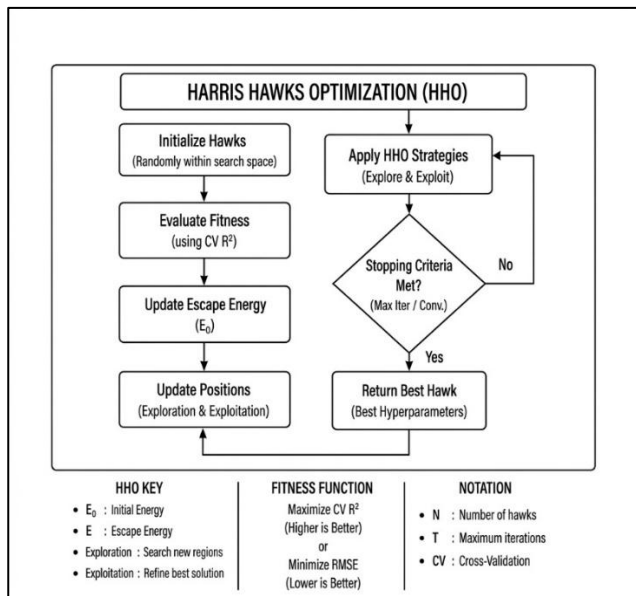


Fig. 8. Harris Hawks Optimization (HHO) working.

Steps to optimize RF model using HHO:

Initialization:

Initialize Lower Bound (LowB) and Upper Bound (UpB) vector: using Eq. (31) and (32)

Initialization of Hawks position matrix: using Eq. (33) and assign values to it using Eq. (34)

Compute Fitness for every hawk using Eq. (35)

Identify the best hawk (Rabbit) as:

$$X_{\text{rabbit}} = \arg \max_i \text{Fitness}(X_i) \quad (58)$$

Calculate the escaping energy that regulates the change from exploration to exploitation using the equation below:

$$E_0 = 2r - 1 \quad (59)$$

where, r is a random value such that $0 \leq r \leq 1$, current escaping energy is expressed as:

$$E = 2E_0 \left(1 - \frac{t}{T}\right) \quad (60)$$

where, t shows the current iteration and T stands for the maximum number of iterations

Exploration phase: When $|E| \geq 1$, the hawk explores the search space as:

$$X(t+1) = X_{\text{random}}(t) - r1|X_{\text{random}}(t) - 2r2X(t)| \quad (61)$$

where, X_{random} is a randomly selected hawk.

Exploitation phase: When $|E| < 1$, hawk exploits the best solution (Rabbit) as :

$$X(t+1) = X_{\text{rabbit}} - E|JX_{\text{rabbit}} - X(t)| \quad (62)$$

where, $J = 2(1 - r)$, and r is a random such that $0 \leq r \leq 1$

Checking new computed value for boundaries using UB and LB as

$$X = \min(\text{UpB}, \max(\text{LowB}, X)) \quad (63)$$

Evaluate Fitness again

Update the Rabbit again

Until the maximum number of iterations is attained, repeat steps 2 through 9.

RF is optimized with best solution as

$$X_{\text{best}} = X_{\text{rabbit}} \quad (64)$$

O. WOA Optimized RF

Seyedali Mirjalili and Andrew Lewis introduced the Whale Optimization method (WOA), a metaheuristic optimization method inspired by nature, in 2016. Only humpback whales exhibit the unusual behavior known as "bubble-net feeding." The whale optimization algorithm (WOA) provides a theoretical description of the spiral bubble-net feeding maneuver. WOA mimics hunting behavior by pursuing the prey with a random or optimum search agent. WOA simulates humpback whale bubble-net attacks using a spiral [35]. Fig. 9 depicts WOA-based RF optimization working.

Steps to optimize RF model using WOA:

Initialization:

Initialize Lower Bound (LowB) and Upper Bound (UpB) vector: using Eq. (31) and (32)

Initialization of Whale position matrix: using Eq. (33) and assign values to it using Eq. (34)

Compute Fitness for every whale using Eq. (35)

Identify best whale as:

$$X^* = \arg \max_i \text{Fitness}(X_i) \quad (65)$$

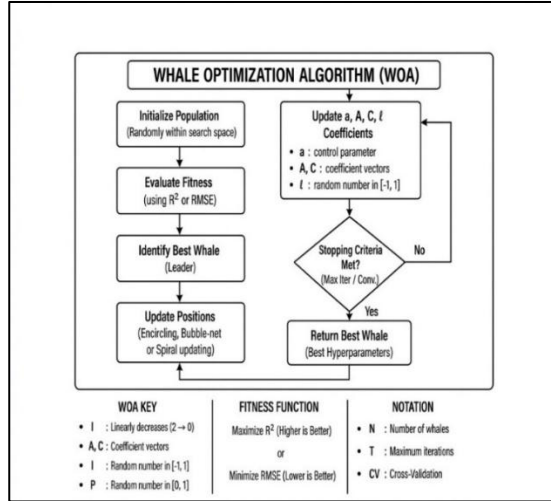


Fig. 9. Whale Optimization Algorithm (WOA) working.

Update control parameter as:

$$a = 2 - \frac{2t}{T} \quad (66)$$

where, t shows the current iteration and T stands for the maximum number of iterations

Compute Coefficient Vectors as:

$$A = 2a r_1 - a \quad (67)$$

$$C = 2r_2 \quad (68)$$

where, A and C are coefficients, where r_1 and r_2 are random values such that $0 \leq r_1, r_2 \leq 1$

Choose the search strategy based on a randomly generated p value between 0 and 1 as follows:

Case 1: If $p < 0.5$ and $|A| < 1$, compute distance to whale and update position as:

$$\text{Distance} = |CX^* - X| \quad (69)$$

$$X(t+1) = X^* - A \text{Distance} \quad (70)$$

Case 2: If $p < 0.5$ and $|A| \geq 1$, compute distance to whale and update position as:

$$D = |CX_{\text{random}} - X| \quad (71)$$

$$X(t+1) = X_{\text{rand}} - A \text{Distance} \quad (72)$$

Case 3: If ≥ 0.5 , compute distance to whale and update position as:

$$\text{Distance}' = |X^* - X| \quad (73)$$

$$X(t+1) = \text{Distance}' e^{bl} \cos(2\pi l) + X^* \quad (74)$$

where, b is the spiral constant, generally taken as 1, l ranges $-1 \geq l \geq 1$

Checking new computed value for boundaries using J and LB as

$$X = \min(\text{UpB}, \max(\text{LowB}, X)) \quad (75)$$

Evaluate Fitness again

Update the best whale as:

$$X^* = X_i \quad (76)$$

Repeat steps 2 to 9 until the maximum number of iterations are reached.

RF is optimized with best solution as

$$X_{\text{best}} = X^* \quad (77)$$

IV. EVALUATION METRICS

The performance of the proposed crop yield prediction framework is examined using R Squared, Mean Absolute Error, Root Mean Square Error, and 5-Fold CV Score.

A. Mean Absolute Error (MAE)

The average absolute difference between actual and predicted data is measured as mean absolute error. It provides results in the target variable's unit. It can be evaluated as:

$$MAE = \frac{1}{N} \sum_{i=1}^N |Y_i - \hat{Y}_i| \quad (78)$$

where, N count of observations is, Y_i is the actual value, $\hat{Y}_i$ is the predicted value.

B. Root Mean Square Error (RMSE)

The square root of Mean Squared Error (MSE) is called Root Mean Squared Error. It improves interpretability by returning the result to the target variable's original unit while maintaining the severe penalty for significant errors. It is evaluated as:

$$RMSE = \sqrt{MSE} \quad (79)$$

$$\text{where, } MSE = \frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y}_i)^2 \quad (80)$$

C. R Squared (R^2)

It is also known as the Coefficient of Determination. R-squared indicates how much of the variance in the target variable can be explained by the regression model. It is judged as:

$$R^2 = 1 - \frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^N (Y_i - \bar{Y}_i)^2} \quad (81)$$

D. Cross-Validation Score (CV)

$$R_k^2 = 1 - \frac{\sum_{i=1}^{N_k} (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{N_k} (Y_i - \bar{Y}_i)^2} \quad (82)$$

$$CV_{\text{Score}} = \frac{1}{k} \sum_{k=1}^5 R_k^2 \quad (83)$$

5-fold CV Score is used in the proposed model as:

$$CV_{\text{Score}} = \frac{R_1^2 + R_2^2 + R_3^2 + R_4^2 + R_5^2}{5} \quad (84)$$

V. RESULTS AND DISCUSSION

All four of the optimized models are found to have outperformed the baseline models. RF Parameters' bounds and best optimized parameters using the PSO optimizer are shown in the following table (Table I):

TABLE I. RF PARAMETERS' BOUNDS AND BEST OPTIMIZED PARAMETERS

RF Parameter	Lower Bound (LB)	Upper Bound (UB)	Final Optimized Value (PSO-RF)
n_estimators	50	500	126
max_depth	5	30	27
min_samples_split	2	20	2
min_samples_leaf	1	10	1
max_features	0.3	1.0	0.951048340383684

RF Optimization resolves black box Interpretation gap: In our model, RF feature selection is used, which prevents dropping domain-critical climate features. RF Optimization resolves the static hyperparameter gap: Our framework addresses this gap by optimizing the RF model using PSO, GWO, WOA, and HHO optimizers, which optimize the RF model based on predefined Lower and Upper Bounds of the RF model's parameters.

PSO-optimized RF performed the best with R^2 Value=98.32. Fig. 10 represents the comparison of models based on R^2 value.

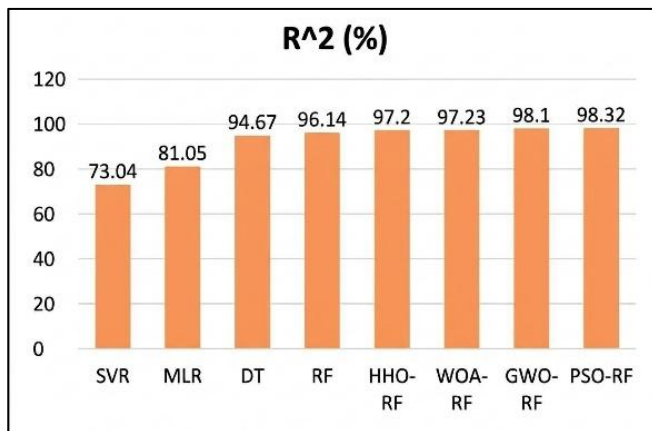


Fig. 10. Comparison of models based on R^2 value.

All four of the optimized models are found to have outperformed the baseline models. And PSO Optimized RF performed the best with CV_Score=97.22. Fig. 11 represents the comparison of models based on the CV_Score value.

All four of the optimized models are found to have outperformed the baseline models. And PSO Optimized RF performed the best with MAE=9.2766. Fig. 12 represents the comparison of models based on MAE value.

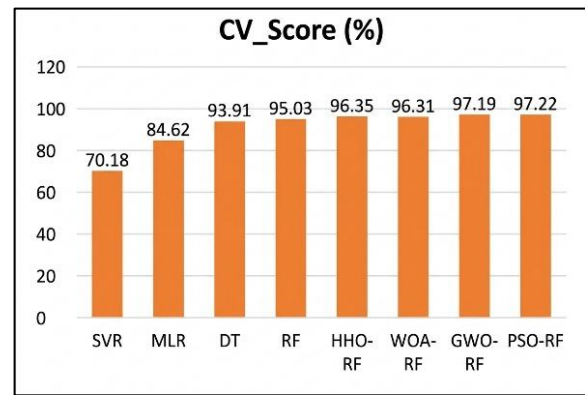


Fig. 11. Comparison of models based on CV_Score value.

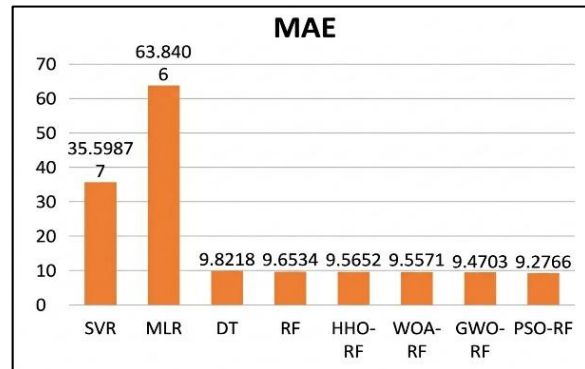


Fig. 12. Comparison of models based on MAE value.

It is observed that RMSE value is quite large. It is due to the dataset, as it contains many samples with 0 or extremely low crop yield. We cannot delete these samples, as it shows that during that period crop yield was very low or the crop was completely wasted, and we want to maintain the originality of the dataset. These yield may be low due to lack of water or weather condition or may be due to pest in the crop or due to weed or due to disaster or may be some other reason. Our algorithms focus on the other samples and predicted low or 0 crop yield values with high error. These samples are few; that is why RMSE and MSE are highly impacted, but have very little impact on R^2 . All four of the optimized models are found to have outperformed the baseline models. And PSO-optimized RF performed the best with RMSE value = 116.1088. Fig. 13 represents the comparison of models based on RMSE value.

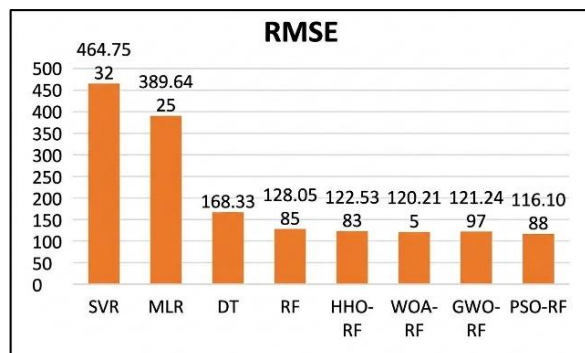


Fig. 13. Comparison of models based on RMSE value.

A. Comparison of Optimized RF Models with Base Models

Tabular comparison of baseline models and Optimized RF models is shown in Table II.

TABLE II. COMPARISON USING PERFORMANCE METRICS

Model	R ² %	CV_Score (%)	MAE	RMSE
SVR	73.04	70.18	35.5987	464.7532
MLR	81.05	84.62	63.8406	389.6425
DT Regression	94.67	93.91	9.8218	168.3300
RF Regression	96.14	95.03	9.6534	128.0585
HHO-RF	97.20	96.35	9.5652	122.5383
WOA-RF	97.23	96.31	9.5571	120.2150
GWO-RF	98.10	97.19	9.4703	121.2497
PSO-RF	98.32	97.22	9.2766	116.1088

VI. CONCLUSION AND FUTURE SCOPE

In this research work, a data analytic framework for crop yield prediction is proposed with a comparison between four baseline models and four optimized RF models. The framework is based on the impact of climate change. The proposed framework mainly comprises of 5 phases: Data input; Data preprocessing; Feature selection based on the RF approach; Train and Test split; After the split all four baseline models are trained and perform crop yield prediction with the computation of four performance metrics: R², MAE, RMSE; 5 Fold CV score is computed; next, out of the four baseline models RF model is chosen as it performed best. RF model is further improved using four nature-inspired models: PSO, GWO, HHO, and WOA. It is observed that all four optimized RF models performed better than the four baseline models. And out of four optimized RF models, PSO optimized performed the best. Therefore, the RF-PSO hybrid model is proposed. The proposed framework's R² (%) value=98.32, MAE value= 9.2766, CV score (%) =97.22, RMSE=116.1088. High RMSE is due to samples with 0 crop yield or very low crop yield. Thus, the proposed framework is suitable for crop yield detection using RF optimization with the PSO optimization technique. In the future, the latest modern metaheuristics can be applied for further improvement in the model.

REFERENCES

- [1] Calvin, K.; Fisher-Vanden, K. Quantifying the indirect impacts of climate on agriculture: An inter-method comparison. *Environ. Res. Lett.* 2017, 12, 115004.
- [2] Prospects by Major Sector (2020, April 10). <http://www.fao.org/3/Y3557E/y3557e08.htm>
- [3] Wheeler, T.; Von Braun, J. Climate change impacts on global food security. *Science* 2013, 341, 508–513. Lesk, C.; Rowhani, P.; Ramankutty, N. Influence of extreme weather disasters on global crop production. *Nature* 2016, 529, 84–87.
- [4] Abbe, C. The physical basis of long-range weather. *Mon. Weather Rev.* 1901, 29, 551–561.
- [5] Modi, A. · Sharma, P. Saraswat, D. Mehta R. Review of crop yield estimation using machine learning and deep learning techniques *Scalable Computing*. 2022; 23(2):59-79
- [6] Van Klompenburg, T. Kassahun, A. · Catal, C. Crop yield prediction using machine learning: a systematic literature review *Comput. Electron. Agric.* 2020; 177(July), 105709
- [7] Samuel, A.L. Some Studies in Machine Learning Using the Game of Checkers. *IBM J. Res. Dev.* 1959, 44, 206–226. [CrossRef]
- [8] Azuaje, F. Review of " data mining : practical machine learning tools and techniques " by witten and frank. January 2006 <https://doi.org/10.1186/1475-925X-5-51> Date: 2014
- [9] Machine Learning in Agriculture: A Review Konstantinos G. Liakos 1 , PatriziaBusato 2, DimitriosMoshou 1,3, Simon Pearson 4 ID and DionysisBochtis 1
- [10] Singh, K. K., Baxla, A. K., Singh, P., & Singh, P. K. (2019). Weather based information on risk management in agriculture. In *Climate change and agriculture in India: Impact and Adaptation* (pp. 207-216). Springer, Cham.
- [11] Ramesh, V., & Kumaresan, P. (2026). Hybrid bio-inspired optimization and Q-learning framework for deep learning crop yield prediction. *Array*, 30, 100982.
- [12] Mohanty N. K., Mallick P.K. and Tripathy H. K., Climate-Sensitive Modeling of Rice Stock in Odisha Using Hybrid Machine Learning and Explainable AI, in *IEEE Access*, vol. 14, pp. 28785-28805, 2026, doi: 10.1109/ACCESS.2026.3665921.
- [13] Chaurasia A., Kaur A., A Hybrid Machine Learning Approach for Accurate Crop Yield Forecasting, *Multimodal Artificial Intelligence in Precision Agriculture*, Edition 1st Edition, 2026, CRC Press
- [14] Manjunath, M. C., & Palayyan, B. P. (2023). An efficient crop yield prediction framework using hybrid machine learning model. *Revue d'Intelligence Artificielle*, 37(4), 1057.
- [15] Dehghanisani, H., Emami, H., Emami, S. et al. A hybrid machine learning approach for estimating the water-use efficiency and yield in agriculture. *Sci Rep* 12, 6728 (2022). <https://doi.org/10.1038/s41598-022-10844-2>
- [16] Juan Cao, Zhao Zhang, Yuchuan Luo, Liangliang Zhang, Jing Zhang, Ziyue Li, Fulu Tao, Wheat yield predictions at a county and field scale with deep learning, machine learning, and google earth engine, *European Journal of Agronomy*, Volume 123, 2021, 126204, ISSN 1161-0301, <https://doi.org/10.1016/j.eja.2020.126204>.
- [17] Saeed Khaki, Lizhi Wang, and Sotirios V Archontoulis. A cnn-rnn framework for crop yield prediction. *Frontiers in Plant Science*, 10:1750, 2020. Zhang L., Zhang J., Li Z., Tao F., Wheat yield predictions at a county and field scale with deep learning, machine learning, and google earth engine, *European Journal of Agronomy*, Volume 123, 2021, 126204, ISSN 1161-0301, <https://doi.org/10.1016/j.eja.2020.126204>.
- [18] Jichong Han, Zhao Zhang, Juan Cao, Yuchuan Luo, Liangliang Zhang, Ziyue Li, and Jing Zhang. Prediction of winter wheat yield based on multi-source data and machine learning in china. *Remote Sensing*, 12(2):236, 2020.
- [19] Jie Sun, Liping Di, Ziheng Sun, Yonglin Shen, and Zulong Lai. Countylevel soybean yield prediction using deep cnn-lstm model. *Sensors*, 19(20):4363, 2019
- [20] PS Maya Gopal and R Bhargavi. A novel approach for efficient crop yield prediction. *Computers and Electronics in Agriculture*, 165:104968, 2019.
- [21] Yousif Y Hilal, Wan Ishak, Azmi Yahya, and Zulfa H Asha'ari. Development of genetic algorithm for optimization of yield models in oil palm production. *Chilean journal of agricultural research*, 78(2):228– 237, 2018
- [22] Alberto González Sánchez, Juan Frausto Solís, Waldo Ojeda Bustamante, et al. Predictive ability of machine learning methods for massive crop yield prediction. *Spanish Journal Of Agricultural Research*, 12(2):313– 328, 2014.
- [23] Ruan, G., Li, X., Yuan, F., Cammarano, D., Ata-UI-Karim, S. T., Liu, X., Tian, Y., Zhu, Y., Cao, W., Cao, Q. (2022). Improving wheat yield prediction integrating proximal sensing and weather data with machine learning. *Computers and Electronics in Agriculture*, 195, 106852.
- [24] Yue, Y., Li, J. H., Fan, L. F., Zhang, L. L., Zhao, P. F., Zhou, Q., ... & Dong, X. H. (2020). Prediction of maize growth stages based on deep learning. *Computers and electronics in agriculture*, 172, 105351.
- [25] Quinlan, J. R. (1986). "Induction of decision trees" (PDF). *Machine Learning*. 1: 81–106. doi:10.1007/BF00116251. S2CID 189902138.

- [26] von Winterfeldt, Detlof; Edwards, Ward (1986). "Decision trees". *Decision Analysis and Behavioral Research*. Cambridge University Press. pp. 63–89. ISBN 0-521-27304-8.
- [27] Yan, Xin (2009), *Linear Regression Analysis: Theory and Computing*, World Scientific, pp. 1–2, ISBN 978-981-283-411-9, archived from the original on 2024-10-04, retrieved 2015-02-07
- [28] Drucker, Harris; Burges, Christ. C.; Kaufman, Linda; Smola, Alexander J.; and Vapnik, Vladimir N. (1997); "Support Vector Regression Machines", in *Advances in Neural Information Processing Systems 9*, NIPS 1996, 155–161, MIT Press.
- [29] Prasad, N. R., Patel, N. R., & Danodia, A. (2021). Crop yield prediction in cotton for regional level using random forest approach. *Spatial Information Research*, 29(2), 195-206.
- [30] Kennedy, J.; Eberhart, R. (1995). "Particle Swarm Optimization". *Proceedings of IEEE International Conference on Neural Networks*. Vol. IV. pp. 1942–1948. doi:10.1109/ICNN.1995.488968.
- [31] Shi, Y.; Eberhart, R.C. (1998). "A modified particle swarm optimizer". *Proceedings of IEEE International Conference on Evolutionary Computation*. pp. 69–73. doi:10.1109/ICEC.1998.699146.
- [32] Mirjalili, Seyedali; Mirjalili, Seyed Mohammad; Lewis, Andrew (2014). "Grey Wolf Optimizer". *Advances in Engineering Software*. 69. ElsevierBV: 46– 61. doi:10.1016/j.advengsoft.2013.12.007. hdl:10072/66188. ISSN 0965-9978. Retrieved 2025-04-02.
- [33] Qiu, Yihui; Yang, Xiaoxiao; Chen, Shuixuan (2024-06-20). "An improved gray wolf optimization algorithm solving to functional optimization and engineering design problems" (PDF). *Scientific Reports*. 14 (1). Springer Science and Business Media LLC: 14190. Bibcode:2024NatSR..1414190Q. doi:10.1038/s41598-024-64526-2. ISSN 2045-2322.
- [34] Ali Asghar Heidari, Seyedali Mirjalili, Hossam Faris, Ibrahim Aljarah, Majdi Mafarja, Huiling Chen, Harris hawks optimization: Algorithm and applications, *Future Generation Computer Systems*, Volume 97, 2019, Pages 849-872, ISSN 0167-739X, <https://doi.org/10.1016/j.future.2019.02.028>
- [35] Seyedali Mirjalili, Andrew Lewis, The Whale Optimization Algorithm, *Advances in Engineering Software*, Volume 95, 2016, Pages 51-67, ISSN 0965-9978, <https://doi.org/10.1016/j.advengsoft.2016.01.008>.
- [36] <https://www.kaggle.com/datasets/akshatgupta7/crop-yield-in-indian-states-dataset>