E-ISSN: 2618-0618
P-ISSN: 2618-060X

www.agronomyjournals.com
2025; SP-8(1): 549-562
Received: 09-11-2024
Accepted: 11-12-2024

Anil Kumar
Assistant Professor, Department of Agronomy, Eklavya University Damoh, Madhya Pradesh, India

Indraveer Singh
Ph.D., Scholar, Farm Machinery and Power Engineering, College of Agricultural Engineering JNKVV, Jabalpur, Madhya Pradesh, India

Mohit Kashyap
Ph.D., Research Scholar, Department of Soil Science, CSK HPKV, Palampur, Himachal Pradesh, India

Ashutosh Kumar
SMS Horticulture Vegetables, KVK Narkatiyaganj, RPCAU, Pusa, Bihar, India

Ningombam Bijaya Devi
Assistant Professor, Department of Botany, G.P. Women's College, Dhanamanjuri University, Imphal, Manipur, India

Shreya Singh
Assistant Professor, Department of Plant Science, School of Advanced Agriculture Sciences and Technology, Chhatrapati Shahu Ji Maharaj University, Kanpur, Uttar Pradesh, India

Shaina Sharma
M.Sc., Department of Agronomy, Chaudhary Sarvan Kumar Krishi Vishwavidyalaya, CSK Himachal Pradesh Agriculture University, Palampur, Himachal Pradesh, India

Rahul Pradhan
Ph.D., Research Scholar in Forestry, Silviculture and Forest Management Department, Institute of Wood Science and Technology, ICFRE-IWST, Bengaluru, Karnataka, India

Corresponding Author:
Indraveer Singh
Ph.D., Scholar, Farm Machinery and Power Engineering, College of Agricultural Engineering JNKVV, Jabalpur, Madhya Pradesh, India

Integration of machine learning and remote sensing in crop yield prediction: A review

Anil Kumar, Indraveer Singh, Mohit Kashyap, Ashutosh Kumar, Ningombam Bijaya Devi, Shreya Singh, Shaina Sharma and Rahul Pradhan

DOI: https://doi.org/10.33545/2618060X.2025.v8.i1Sh.2496

Abstract
Increasing demand for accurate crop yield predictions in agriculture has been fueled by technological innovation. Machine Learning (ML) and Remote Sensing (RS) have become leading tools for precision and scalability in predictions. This review discusses the present state of the integration of ML and RS, bringing out methodologies, datasets, applications, and challenges in predicting crop yields. It has thus proven to show considerable promise for agricultural decision-making, resource use optimization, and improvement in food security through synergies between ML algorithms and RS data. The future trends and potential advancement are also discussed below.

Keywords: Crop yield prediction, machine learning, remote sensing, precision agriculture, agricultural technology, data analytics, satellite imagery, AI in agriculture

1. Introduction

Productivity and sustainability in agriculture are vital in the context of increasing population and climate variability, for food security challenges worldwide (Ahmed *et al.*, 2023) [2]. Precise crop yield estimation can ensure effective planning, optimal resource use, and risk management against the variability of climate, extreme weather conditions, and market volatility. Combining ML with RS technology presents a powerful and innovative way to meet these challenges (Wang *et al.*, 2024) [115]. This integration supports precision agriculture, enhances decision-making, and ultimately contributes to sustainable insights of RS data (Sabir *et al.*, 2024) [97].

Crop yield prediction is an integral part of contemporary agriculture, aiming to overcome the major challenges of food security, optimal resource use, and climate change (Sharma *et al.*, 2024) [105]. Traditional yield-prediction techniques, mainly dependent on manual observation and statistical modeling, fail to capture the full complexity and variability of agricultural systems (Jin *et al.*, 2018) [47]. The machine learning (ML) approach that integrates remote sensing (RS) technologies has turned out to be a game changer, utilizing big data to promote accuracy and scalability (Singh *et al.*, 2021) [109].

There are significant areas of concentration in the continuing global effort towards food security. Given the rapidly increasing global population and the challenges caused by climate change and extreme weather events, demand for innovative solutions in agriculture has never been more pressing. The integration of accurate crop yield predictions can play a pivotal role in ensuring effective agricultural planning, resource optimization, and risk mitigation strategies, especially in times of climate variability and market instability (Akintuyi, O. B. 2024) [3].

In this context, Machine Learning (ML) and Remote Sensing (RS) technologies offer a promising solution to meet these challenges. By taking the strong predictive ability of ML algorithms and combining them with the insight's RS data contains on the spatial and time dimensions, sophisticated models can be developed which not only enhance the accuracy of crop yield forecasts but also support the optimization of farming practices (Polwaththa *et al.*, 2024) [92].

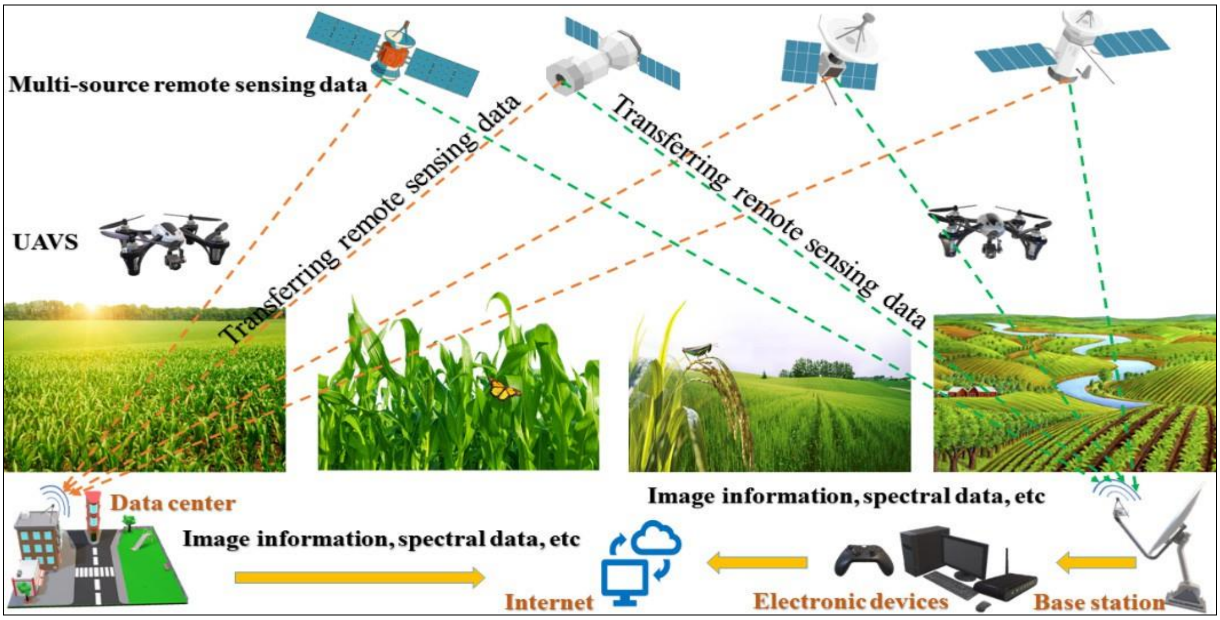


Fig 1: Use of remote sensing for crop monitoring

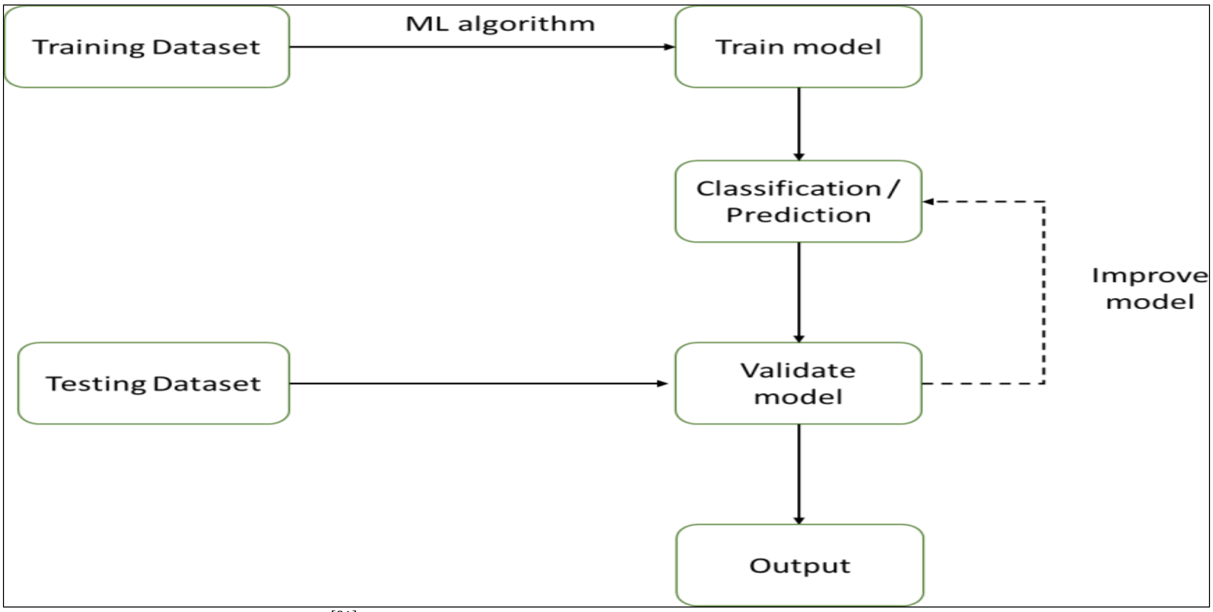
The innovative integration will allow the implementation of precision agriculture and better decision-making processes while contributing to more sustainable agricultural systems, and hence food security, on a large global scale (Bhat SA, Huang NF, (2021) [15].

This review reviews the recent trend in ML and RS for crop yield prediction and points out improvements in data acquisition via satellites and UAVs, feature extraction, and model development. Key machine learning algorithms used for the processing of RS data include random forests, support vector machines, deep learning, and ensemble methods in multispectral, hyperspectral, and LiDAR imagery (El-Omairi, and El Garouani, 2023) [30]. In addition, the review deals with challenges regarding data quality, spatial resolution, model interpretability, and transferability across regions. By synthesizing recent studies and identifying gaps in research, this work sheds light on the possibility of revolutionizing sustainable agriculture and decision-making processes through ML and RS (Han *et al*., 2023) [42].

2. Machine Learning in Agriculture

Machine Learning is the next generation of changes sweeping through almost all sectors; it has already impacted agriculture greatly. The high capabilities of the algorithms to deal with and analyze complex data have highly boosted the use of crop yield predictions and much better farming methods (Elbasi *et al*., 2023) [29]. By enabling farmers to gain insights from diverse data sources such as weather patterns, soil conditions, crop health indicators, and management practices, ML technologies have opened up new avenues for increasing agricultural productivity and ensuring sustainability (Karunathilake *et al*., 2023) [52].

With the aid of advanced data analysis techniques, ML models can detect subtle interrelations among variables, thereby allowing the prediction of future outcomes, optimization of agricultural practices, and early detection of potential problems such as pest outbreaks, diseases, and weather anomalies. This is why the incorporation of ML into agricultural systems is of great promise to enhance resource use efficiency, crop yields, and environmental impact (Alahmad *et al*., 2023) [4].



Source: Pokhariyal *et al*., 2023 [91]

Fig 2: Essential architecture for Machine learning

2.1 Overview of Machine Learning Algorithms

The core of predictive analytics in agriculture is the ML algorithms. These algorithms differ in their approaches and methodologies to solve the vast spectrum of agricultural challenges-from predicting crop yields to detecting pests and diseases (Alzubi *et al*., 2018) [6]. The choice of algorithm depends on the specific task at hand and the nature of the available data. Below, we explore different categories of ML algorithms and their applications in agriculture.

Regression Models

Regression models are very common for predicting continuous outcomes; hence, the application can be optimal for deriving estimates of crop yields from environmental and agricultural data (Fahrmeir *et al*., 2013) [31]. These are generally used when there is a clear, measurable relationship between the dependent variable, like crop yield, and the independent variables, such as weather conditions and soil attributes. Among the common regression models are the following.

- **Linear Regression:** This model is adopted for the prediction of crop yield based on straightforward relations with variables such as temperature, rainfall, and soil quality (Conradt T, 2022) [23]. It supposes a linear relationship between the input variables and the output.
- **Support Vector Regression:** A more advanced method, it handles complex, non-linear relations in the data. SVR is highly useful and effective when the relation of input features with the outputs is not straightforward (Zhou *et al*., 2024) [125].
- **Ridge and Lasso Regression:** These models introduce regularization terms that prevent overfitting, particularly when datasets contain many predictors (Santos *et al*., 2022) [100]. Ridge regression can handle multi-collinearity; otherwise, Lasso regression can be used for feature selection because some coefficients get forced to zero.

Tree-Based Models

Tree-based models are effective at handling nonlinear and high-dimensional data, common in agriculture (Nzekwe *et al*., 2024) [84]. Tree-based models operate on the basis of recursive partitioning of the data using feature importance (Badshah *et al*., 2024) [12]. The approach ensures complex interactions between variables are managed well. Major tree-based models are:

- **Decision Trees:** Decision trees are a way of breaking down data into smaller, more manageable parts by making decisions based on the most significant features (Charbuty, & Abdulazeez, 2021) [19]. They are interpretable, making them valuable for understanding which factors are most important for predicting outcomes like crop yields.
- **Random Forests:** Random Forest is an ensemble method that builds up multiple decision trees and aggregates their results in order to offer a better accuracy and avoid overfitting (Zhou ZH, 2025) [126]. This seems to be especially useful for enhancing robustness and performance of predictions.
- **XGBoost:** An advanced gradient-boosting algorithm that combines the predictions of multiple decision trees. XGBoost is known for its speed, scalability, and high performance, making it particularly effective when working with large datasets common in agricultural research (Elavarasan and Vincent, 2020) [28].

Deep Learning Models

Deep learning models are powerful tools to handle large, unstructured datasets such as satellite images, time series data, and sensor readings (Moskolaï *et al*., 2021) [79]. These models, inspired by the structure of the human brain, consist of multiple layers of interconnected nodes (neurons) that learn hierarchical representations of the input data (Cravero *et al*., 2022) [24]. Notable deep learning models include.

- **Convolutional Neural Networks:** CNNs are very efficient at analyzing spatial data, like images, and can be used to classify images for assessing crop health by using satellite or drone imagery (Bouguettaya *et al*., 2022) [17]. These models extract relevant features automatically from images, which makes them highly valuable for monitoring crop conditions and detecting diseases or pests (Peng *et al*., 2024) [90].
- **Recurrent Neural Networks (RNNs):** RNNs are particularly designed for sequential data, such as time series data that could represent weather patterns, crop growth stages, or sensor readings over time (Khaki *et al*., 2020). They are especially useful in forecasting trends and predicting future crop yields based on historical data (Kamilaris *et al*., 2018) [50].
- **Transformers:** The most complex architecture developed lately that is increasingly gaining acceptance nowadays (Giedion, 2009) [34]. These excel at learning long-range dependencies within data and hence prove quite beneficial when using disparate data streams together such as overlaying weather forecast information over satellite images, solving the problems associated with agriculture under multiple models.

Clustering and Unsupervised Learning

Unsupervised learning techniques, such as clustering, are used to identify patterns and groupings within data without requiring labeled outcomes (Mehta *et al*., 2015) [75]. This proves very helpful in the discovery of huge datasets or when trying to uncover hidden relationships (Badapanda *et al*., 2022) [11]. Some of the popular clustering algorithms include:

- **K-Means Clustering:** A widely used algorithm that divides data into k clusters based on similarities in features (Golubovic *et al*., 2019) [37]. For example, k-means can be used to segment agricultural regions based on soil types or vegetation indices, enabling targeted interventions for different regions (Pascucci *et al*., 2018) [87].
- **Principal Component Analysis**: PCA is another technique for dimension reduction that helps make complex datasets much easier to view and understand by bringing out the important features or principal components of those data sets (Greenacre *et al*., 2022) [38].

Hybrid Models

Hybrid models are ones that integrate various machine learning algorithms together in an attempt to draw strength from every individual model used (Krasnopolsky *et al*., 2006) [59]. These models combine the interpretability of simpler algorithms with the power of complex ones, offering practical solutions to intricate agricultural problems. For example, combining Random Forests with Neural Networks offers the best of both worlds: interpretability as well as prediction accuracy, which can serve to better enable farmers and researchers to understand the factors driving the predictions and at the same time benefit from robust, data-driven insights (Aria *et al*., 2021) [9].

In summary, the integration of different machine learning algorithms into agricultural systems offers great potential to revolutionize farming practices, enhance productivity, and contribute to sustainability (Sharma *et al*., 2020) [106]. By harnessing the power of advanced models, farmers can make more informed decisions, optimize resources, and ultimately improve global food security.

2.2 Applications in Crop Yield Prediction

ML has revolutionized crop yield prediction as it provides significant improvements in terms of the accuracy and reliability of the prediction made (Rashid *et al*., 2021) [96]. Previously, crop yield prediction used only simple statistical models that involved fewer factors. However, with ML, many variables related to climate conditions, soil characteristics, and crop health indicators can be analyzed together to create more accurate and dynamic yield predictions (Chlingaryan *et al*., 2018) [21]. The features of ML in the sense that it can work on very large datasets and recognize patterns in them offer numerous applications in crop yield prediction that fall under a few main clusters.

Data Integration

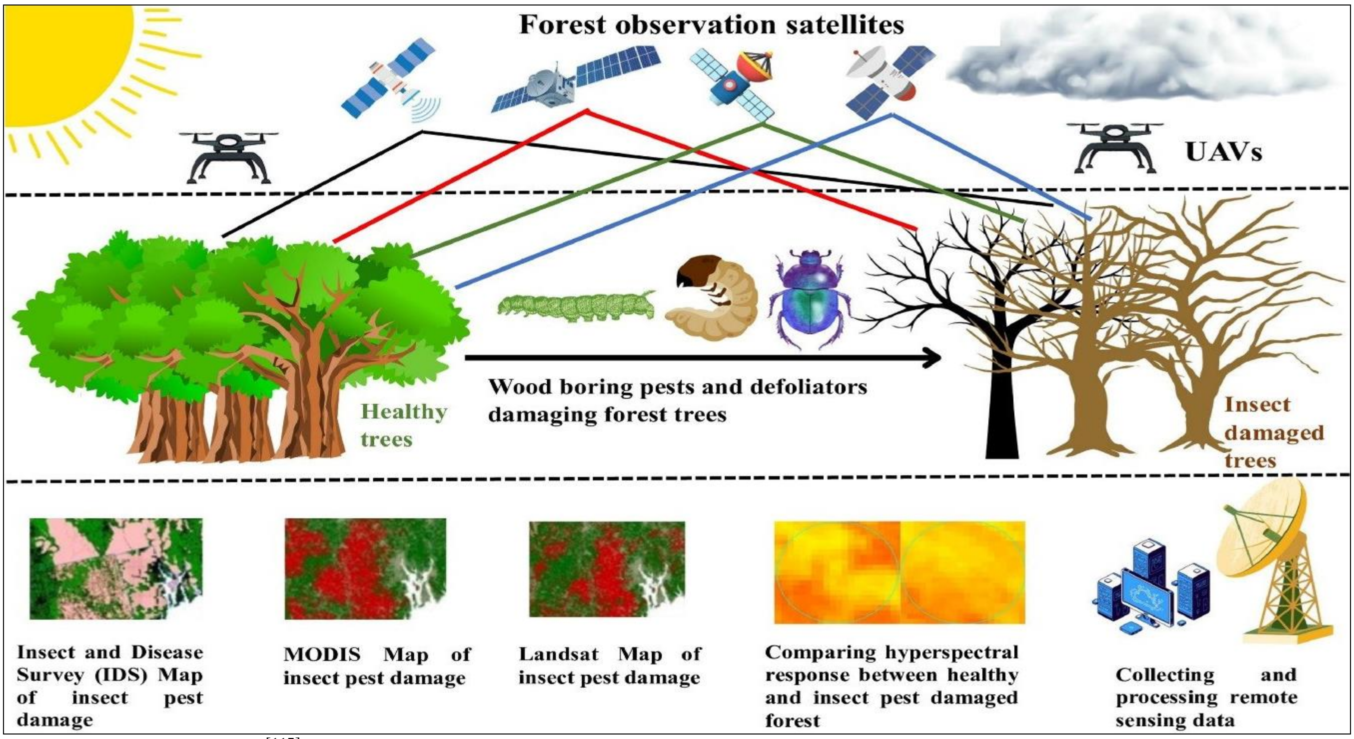
Crop yield prediction is one of the most useful applications of ML. Nowadays, modern agriculture systems produce highly diverse data originating from multiple channels of satellite imagery, IoT sensors, historical yield records, and live weather patterns (Saiz-Rubio, and Rovira-Más, 2020) [98]. ML models particularly thrive in condensing this vast and voluminous data toward obtaining a holistic picture of the environment. By recognizing interdependencies between different variables, ML models can make more accurate predictions about crop yields (Chlingaryan *et al*., 2018) [21].

For instance, IoT enabled sensors placed in the field can collect real time soil moisture, temperature, and nutrient levels. Thus combining the above real-time data with historical yield data and weather forecasts, can enable more accurate models in predicting future crop yields. The integration also supports adaptive management practices by allowing proactive farmer responses to environmental changes.

Integration of Remote Sensing

Remote sensing is considered a very important method for crop yield prediction, as it delivers valuable information related to vegetation health, growth patterns, and environmental conditions (Karthikeyan *et al*., 2020) [51]. Satellite and drone-based remote sensing technologies offer very high-resolution images that include vegetation indices (such as NDVI), spectral reflectance, and temperature data (Bansod *et al*., 2017) [13]. Such parameters are extremely important inputs into ML algorithms targeted at high accuracy crop yield estimates (Paudel *et al*., 2021) [89].

An example of this integration is the use of MODIS satellite data, which offers global coverage and high-frequency data on vegetation health and temperature (Yang, *et al*., 2020) [122]. By applying Random Forest models to MODIS data, researchers can predict crop yields, such as corn, across vast geographic areas. Remote sensing data can also be combined with ground-level data, such as soil characteristics and farm management practices, to improve yield predictions (Argento *et al*., 2021) [8].



Source: Wang *et al*., 2024) [115]

Fig 3: The process of detecting insect infestation and evaluating forest health was realized based on IDS, MODIS and Landsat-8 mapping, along with hyper spectral response technology

Scenario Analysis

ML models can be used not only for predicting current crop yields but also to simulate potential future scenarios. They can assess how changes in agronomic practices, such as planting dates, irrigation schedules, and fertilizer application rates, may affect crop yield. By examining a variety of scenarios, ML can give decision-makers insights into how to optimize farming practices for maximum productivity under varying conditions (Sharma *et al*., 2020) [106].

For instance, ensemble learning methods have been utilized to predict the yields of wheat under various scenarios of climate change (Iqbal *et al*., 2024) [45]. Such models consider how shifts in temperature regimes, alterations in rainfall patterns, and changes in other climate components might impact crop growth and overall production (Eddamiri *et al*., 2024) [27]. Simulated scenarios help farmers and agricultural planners design more robust strategies for coping with climate-related uncertainties.

Automation in Agriculture

One of the most intriguing agricultural technologies used is ML-based predictive models. Predictive models integrate into farm management systems, then automate decision-making so as to make the utmost time to irrigate, fertilize, or control pests (Mohyuddin *et al*., 2024) [78]. Using data in real time from sensors and short-range forecasts, among other sources, ML models can make timely recommendations to farmers, thus reducing waste and improving productivity (Imade *et al*., 2024) [44].

For example, an ML-controlled irrigation system would suggest changes in water usage due to current levels of soil moisture, weather forecasts, and crop-specific water requirements. In the same way, ML can calculate the optimal level of fertilizer or pesticide to use so that every resource is efficiently utilized while crops can grow optimally.

Case Studies

The practical applications of ML in crop yield prediction are becoming increasingly evident, as several case studies demonstrate the effectiveness of ML techniques in various crops. Below are some examples of successful ML applications in crop yield prediction.

Maize Yield Prediction

These Random Forest models were implemented to predict the yield of maize by including information from diverse data sources that relate to the variables of interest like soil properties, climatic variables, and even vegetation indices like NDVI (Ngie, and Ahmed, 2018) [81]. The analysis results in proper estimations for local as well as regional scales, while this form of predictive modeling gives more accurate direction to manage their resources appropriately-water as well as fertilizers-and in consequence, their yield of maize gets improved (Sakamoto *et al*., 2020) [99].

Wheat Yield Prediction

CNNs have been particularly useful in predicting wheat yields based on time-series weather data and crop phenological stages (Wang *et al*., 2023) [114]. Deep learning models are better than traditional regression techniques because they capture complex, non-linear relationships between environmental variables and crop growth. The ability of CNNs to process large amounts of data quickly and efficiently makes them a valuable tool for predicting wheat yields, especially in large-scale farming operations (Chlingaryan *et al*., 2018) [21].

Rice Yield Prediction

Support Vector Machines (SVMs) represent a popular algorithm in supervised learning, and various studies have achieved successful rice-yield predictions that integrate agroclimatic datasets with those of pest infestations (Rashid *et al*., 2021) [96]. Indeed, the inherent ability of SVMs to explore patterns and determine predictions, even based on noisy data or data deficits, can serve as a benefit in such adaptability for different ML models-pest infestation challenges that help ensure reliable forecasts of crop yield (Gandhi *et al*., 2016) [32].

Soybean, Cotton, and Other Crops

Thermal RS data-driven mapping of water stress in soybean fields using regression models. Cotton yield prediction using UAV imagery and ensemble ML techniques.

This paradigm shift in crop yield prediction by Machine Learning not only transforms the traditional approach to agriculture but also empowers stakeholders with critical tools to tackle key challenges related to future food security, climate change, and resource management (Leo *et al*., 2021) [65].

Broader Implications for Agricultural Sustainability

It is not just the application of ML in crop yield prediction but it also represents a paradigm shift for the approach of agriculture. More precision in the prediction of crop yields through ML models enables farmers to make data-driven decisions that help in increasing productivity and reducing waste and environmental degradation (Titirmare *et al*., 2024) [112]. Furthermore, the application of ML provides recommendations tailored to specific farms based on their unique features.

Further, yield prediction models through machine learning promote sustainability of agriculture through the encouragement of resource optimization and support to better climate-resilient strategies (Patil, 2024) [88]. By forecasting likely challenges and suggesting solutions adapted for specific conditions, machine learning assists farmers in preparation against uncertain future conditions, reducing climate change risks related to market volatilities and resource shortages (Amiri *et al*., 2024) [7]. This basically marks a critical aspect of the field of agriculture's evolution and highlights how machine learning contributes to crop yield prediction, allowing for more sustainable and efficient and resilient farming practices that are fundamental in providing food globally. In summary, it is with this adoption of Machine Learning in agriculture that it creates a data-driven transformation reshaping the approaches of farmers, researchers, and policymakers in tackling agricultural productivity and sustainability (Mohamed, 2023) [77].

Table 1: Different techniques used for prediction yield/production

Paper	Dataset	Algorithm	Observation
Pallavi *et al*. 2021 "Crop yield forecasting using data mining".	Agricultural dataset Crop, area, yield, temperature, Humidity	XGBoost, Logistic Regression, Random Forest, KNN	The best classifier with a 67.8 percent accuracy rate is Random Forest.
Mohsen *et al*. 2021 "Crop yield prediction in US corn belt."	A dataset from 1984 to 2018 is used to forecast corn output. Weather data, soil data, corn yield data.	Linear Regression, LASSO, Light GBM, Random Forest, XGBoost	This paper proposed a hybrid simulation-machine learning approach that provided improved county-scale crop yield prediction. This study demonstrated that on average, introducing APSIM variables into machine learning models and utilizing them as inputs to a prediction task can decrease the prediction error measure by RMSE between 7 and 20%.
"Corn and Soybean yield prediction using Deep Learning" (S. Khaki *et al*.,	The yield performance dataset includes the observed county-level average yield for corn and soybean	Random Forest, DFNN, CNN, RT, LASSO, Ridge	Paper proposed a new convolutional neural network model called YieldNet. The goal of this paper is to predict the average yield per unit area of two crops

2021) [55]	between 2004 and 2018 across 1132 counties for corn and 1076 for soybean within 13 states of the US Corn Belt and corn satellite data used MODIS, cropland data layer		before harvest based on the sequence of images taken by satellite. The transfer learning approach maximizes prediction accuracy.
"An ensemble algorithm for crop yield predictions".	From different databases, 7 features were retrieved in 28242 instances of finalized parameters of rainfall, temperature, climatic conditions, crop-type	AdaBoost Regressor, Gradient Boosting Regressor, Tree Regressor, Random Forest Regressor, KNN, Decision Bagging Classifier	Decision tree regressor and Ada Boosting algorithm combinations gives the best results. Decision Tree Regressor has the highest R2 score of 95%. After analysis, AdaBoost and Decision Tree Regressor give the most accepted outcomes with 95.7% accuracy.
"Artificial Neural Network-Based Crop Yield Prediction using NDVI, SPI, VCI feature vector" (P. Tiwari *et al*., 2019) [113]	Geospatial data used various feature vectors of SPI, NDVI and VCI. The Indian state of Madhya Pradesh is the chosen region.	"Error Back Propagation Neural Network" (EBPNN), Spiking Neural Network (SNN)	Numerous assessment factors have been overcome by comparing the proposed model to earlier researchers' studies. Compared to SNN, the proposed approach lowered relative error by 33% and RMSE by 31.8%, improving accuracy overall.
"Environment Monitoring System for agriculture application using IoT and predicting crop yield using various data mining technique"	Datasets are collected from the sensors used at the observation site. Humidity, Temperature, Moisture, Soil, Ph value, CO_2%, Monthly rainfall. Datasets taken from records of Indian Government area and values of crop yield are also taken.	Random Forest, KNN, SVM.	Eleven thousand record values altogether information is used for observation. From random forest classification, it is concluded that it is an efficient tool for prediction on a narrow range of records, while KNN is recommended for a large range of statistical values.
"Supervised Machine Learning approach for crop yield prediction in the agriculture sector".	Rainfall, Temperature, crop name, Ph, and Humidity data are obtained from the Kaggle website.	Random Forest Classifier, Decision Tree	The best crop production is predicted using the Random Forest method.
"A CNN-RNN framework for crop yield prediction" (S Khaki, *et al*., 2020) [55]	Yield dataset for corn and soybean, weather data (Minimum/Maximum, Temperature, solar radiation, Precipitation, Vapor Pressure), data on soil from the whole corn belt of US for 2016, 2017, 2018 years historical data.	LASSO, Deep Fully Connected Neural Networks (DFNN), Random Forest, Recurrent Neural Networks (RNN), CNN	The proposed model outperformed other methods like LASSO, Random Forest, and DFNN. The CNN-RNN model can generalize and provide predictions for untested places. The suggested study accurately predicted corn and soybean yields over the whole corn belt of the United States. Future projects involving yield prediction may make use of the presented model. The CNN-RNN model underwent feature selection using the back propagation approach.
Nishant *et al*. 2020 [83] "Crop Yield Prediction based on Indian Agriculture using Machine Learning".	Information was obtained from an Indian government resource. District, state, season, crop, production, year, and area with 2.5 lakh observations.	Regression techniques like-LASSO, ENET, Kernel Ridge	Root mean squared error is the performance statistic utilized for this assignment. ENET around 4%, LASSO was 2%, and Kernel Ridge after stacking was much less than 1%A web application is designed for user/farmer to enter details to get predictions.
"Wheat crop yield prediction using new Activation Functions in Neural Networks" (SH Bhojani, *et al*., 2020) [16]	Datasets from 1990-1991 to 2016-2017 yield datasets are acquired from the directorate of agriculture in Gandhinagar, Gujarat, for analysis and research purposes. Weather data from the department of agro meteorology	Neural Networks technique of Data Mining, Multi-layer perceptron (MLP)	Compared the outcomes of several activation functions and suggested three new, straightforward activation features-DharaSig, DharaSigm, and SHBSIG-for improving general effectiveness of neural networks and the precision of the findings. Experiments indicate that newly created activation features offer better outcomes than the 'sigmoid' default ANN activation function for agriculture datasets.
"Comparative Evaluation of Neural Networks in Crop Yield Prediction of Paddy and Sugarcane crop" (K. Krupavath *et al*., 2022) [60]	Crop-sensitive parameters extracted from LANDSAT8 OLI imageries. Data on paddy and sugarcane crops have been taken through regional sensing at regional levels.	Feed Forward Back Propagating Neural Network" (FFBPNN)	The results showed that employing remote sensing photos and Neural Network models is very effective; Paddy shows the higher value of Mean Relative Error to be 6.166% and minimum relative error as-0.133%. For Paddy higher, an MAE of 0.178 was recorded.
Manivasagam *et al*. 2022 "An efficient crop yield prediction using machine learning".	The dataset contains production details, soil and environment parameters, rainfall, temperature, Humidity, ph.	Logistic Regression, KNN, Random Forest Classifier	Using ML methods to predict crops and also yield. Comparing the accuracy of three models, Decision tree (90%), KNN (85%), and Random Forest get more accuracy (95%).
"Crop yield prediction techniques using machine learning algorithms".	Nitrogen, Phosphorus, Ph level, Humidity, Temperature	Random Forest, Decision Tree, Logistic Regression, Support Vector Machine, Naïve Bayes	The machine learning methods work on data training, resulting in Naïve Bayes to get optimal accuracy that helps suggest farming practices.
"A comparison between	Cropland data layer: Pixels extracted and registered on the CDL	Deep Neural Networks, ANN,	This article examined several AI algorithms to create a reliable crop yield forecast model. Using
Major Artificial Intelligence models for Crop yield prediction; Case study of the Midwestern United States,	Map as soybean and corn fields. Meteorological data-TMEAN, TMIN, TMAX, PPT. Satellite image data.	Multivariate Adaptive Regression Splines(MARS), Extremely Randomized	A prediction error of around 7.6 percent and 7.8 percent for corn and soybean, the DNN model with the JA (July-August) database beat the other five AI models. For corn and soybean, the model indicated correlation

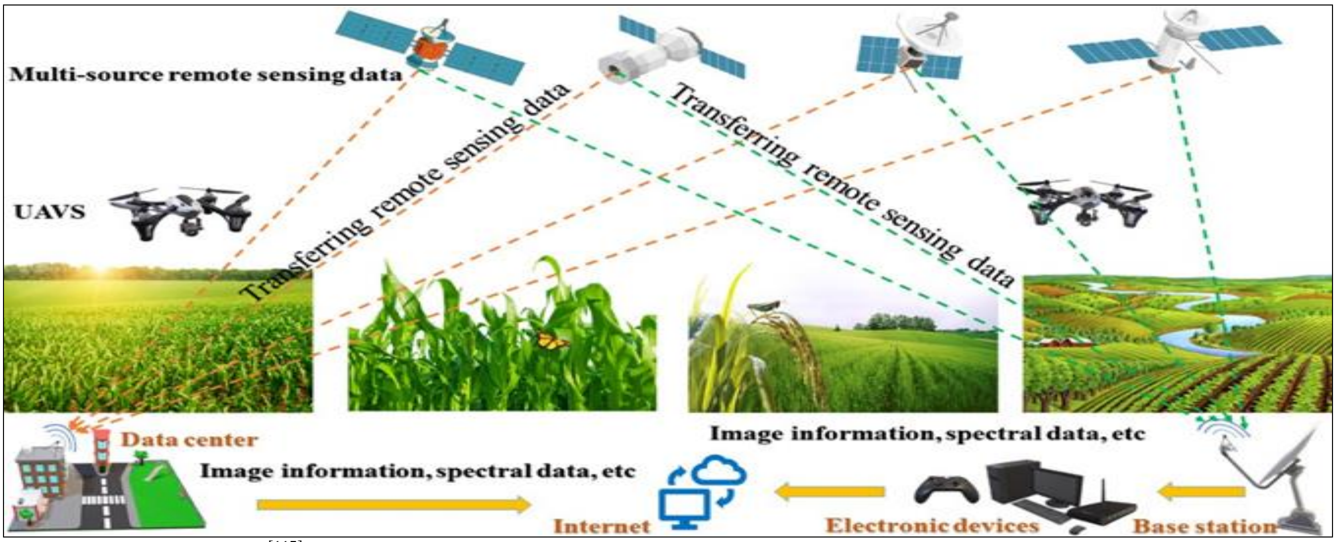
2006-2015".	Statistics on Crop yield-soybean and Corn US Agriculture Department Hydrological data-Soil Moisture	Trees, Random Forest, Support Vector Machine (SVM)	coefficients of 0.945 and 0.901, respectively. This demonstrates how precisely the DNN with the JA model can anticipate corn and soybean production. The Optimized DNN version created for this research may also be used for other crops and in various geographies.
"Smart Farming Systems; crop yield prediction using regression techniques".	The yield and weather databases included four hundred twenty-three observations of corn yield and monthly averages for temperature and precipitation. Rainfall, temperature, yield, and Humidity are all factors that affect the yield of crops.	Random Forest, SVM Regression, Multivariate Polynomial Regression,	Crop yield was used to compare the algorithms based on predicted yield, MAE, RMSE, and R squared and made. SVM outperforms the other two models across the board in all three metrics. Values of Median Absolute Error: 1.58, MAE: 3.57, and RMSE: 5.48, and the max R-squared value for SVM was obtained to be 0.968.
"Yield predict; a crop yield prediction framework for smart farms."	Data is collected from Open Government Data Platform India and also the agricultural site of Rajasthan. Data is divided into two datasets-Kharif crops and Rabi crops are based on the rainfall pattern of India. Additional monitored attributes are District name, State name, Season crop, Crop year, Production, Area, Nitrogen, Seasonal rainfall, Potassium, Phosphorus, Solidic soil (ha), and Saline soil (ha),	K-Nearest Neighbors (KNN), Linear Regressor, Support Vector Regression, Ada Boost Regressor, Extreme Gradient Boosting (XGBoost), Decision Trees, Gradient Boosting Regressor, Random Forest Regressor, Light GBM Regressor.	The Gradient Boosting Regressor model has the greatest R2 value (0.616) compared to other models and performed the best for the Rabi crops dataset. This model's RMSE value is 0.482, which is very low. The XGBoost Regressor, which has an R2 value of 0.572 and an RMSE of 0.37, is the model that performs the best for datasets on Kharif crops.
"Crop yield prediction and efficient use of fertilizers" (S. Bhanumathi *et al*., 2019) [14]	District, State, Area, Crop, Production, Season. This data is applied for training the model and predicting production. Another dataset is used to predict the amount of fertilizer used; input parameters are nitrogen quantity, phosphorus & potassium.	Random Forest algorithm, Back Propagation algorithm	While comparing the two models, the error rate was low with Random Forest as compared to the backpropagation.
"Recommendation system for crop identification and pest control technique in agriculture" (S Bhanumathi *et al*., 2019) [14]	Crop includes Bajra, Cashew nut, Chickpea, Jowar, Cotton, Wheat, Jute, Tea, Sugarcane, Rice, Ragi, and Pulse. Considered Attributes are temperature, average rainfall, ph, and soil color.	SVM, Decision Tree, Logistic Regression	The SVM algorithm gives the highest accuracy of 89.66%. A recommendation system is built where the training set is categorized once input is collected through a form.
"Sugarcane yield grade prediction using Random Forest and Gradient Boosting Tree techniques" (P. Charoen *et al*., 2018) [20]	The study uses data on sugarcane output supplied by a Thailand-based sugar mill. Data are obtained by farmer plots (cane type/class, type of soil, area, fertilizer) and actual yield which the farmers deliver to the mills.	Gradient Boosting; Random Forest Classification.	Accuracies of the two models are-71.83% (Random Forest), and 71.64% (Gradient Boost). Comparisons are made of predictions of the RF-based method and GBT-based method with two different non-machine learning baselines.
"Crop Prediction on the Region Belts of India; A Naïve Bayes Map Reduce Precision Agricultural Model" (R. Priya *et al*., 2018) [94]	The data was collected from Krishi Vigyan Kendre of Telangana State of India. Information is gathered from various sources, including sensor-recorded field data, satellite images, reports on irrigation, crop data, and meteorological information. Temperature, Humidity, wind direction, soil moisture, speed, diffusion rate, radiation, and rainfall are all parameters.	Naïve Bayes Classifier	The proposed paper discusses a system for recommending crops. The proposed system uses MapReduce functionality for evaluation with Naïve Bayes. The model works in two stages-the first suggests which crop should be grown. The second suggests the best sowing month so that growth can be maximized.

Source: Kulyal M & Saxena P, 2022) [62]

3. Remote Sensing in Agriculture

RS technologies are now an inevitable tool in modern agriculture. These have provision for exact, large-scale data for scanning or monitoring and managing agricultural landscapes (Weiss *et al*., 2020) [116]. With the use of satellites, drones, and aerial imagery, it becomes easy to collect real-time data and history for critical decision-making in crop production and yield estimation (Jewiss *et al*., 2020) [46].

Remote Sensing (RS) is revolutionizing modern agriculture by making it possible to monitor large areas of agricultural landscapes at a very fine scale (Ozdogan *et al*., 2010) [85]. In the collection of real-time and historic data, RS through satellites, drones, and aerial imagery assists in informed decision-making on crop yield, environmental monitoring, and yield estimation (Yadav *et al*., 2023) [120].



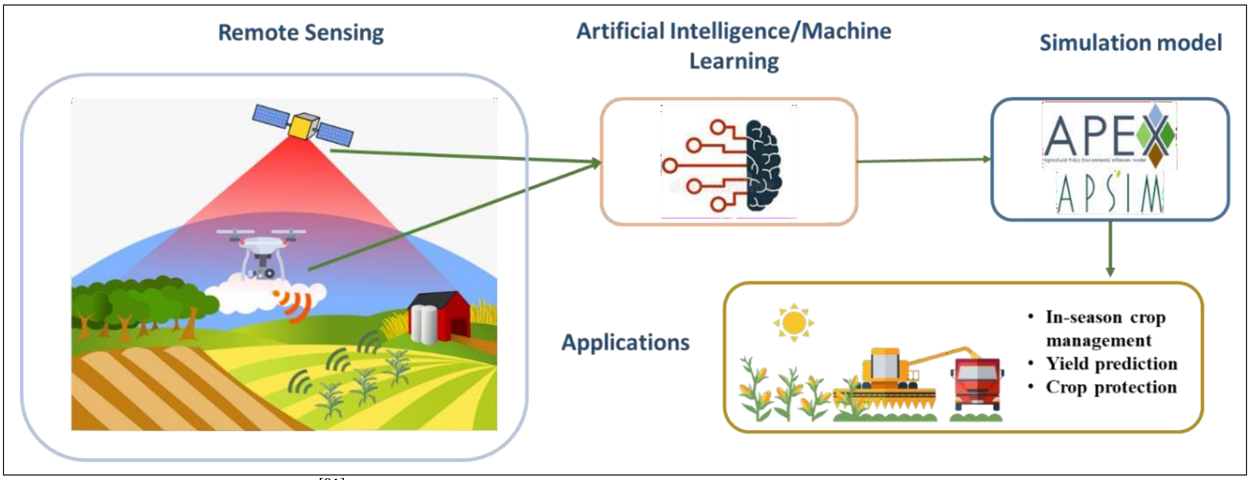
Source: Wang *et al*., 2024 [115]

Fig 4: The comprehensive application framework of different remote sensing satellites in precision agriculture

3.1 Role of Remote Sensing

One important role of RS in agriculture lies in providing detailed spatial and temporal data regarding a wide array of environmental and crop health parameters (Singh *et al*., 2020) [108]. The use of RS allows monitoring to be carried out on many of the factors affecting agriculture directly (Liaghat, and Balasundram, 2010) [68].

- **Monitoring of Vegetation:** The plant health can be tracked by using remote sensing based on the spectral signature of crops (Houborg *et al*., 2015) [43]. The NDVI or Normalized Difference Vegetation Index helps determine plant greenness, vigor, and stress. These values are very crucial in finding out the first sign of disease or water and nutrient deficiency and thereby correcting these measures before problems go out of control (Lawley *et al*., 2016) [64].
- **Soil Analysis:** RS tools can detect soil properties like moisture content, texture and organic matter. Satellite-based soil moisture sensors provide important data for judicious irrigation management. The data are used in conserving water and efficiently improving crop yield (Lobell *et al*., 2015) [69]. It has helped farmers track soil properties within large areas to sustainably manage soil health.
- **Land Use Mapping:** Monitoring and mapping of agricultural region land-use patterns can only be effectively realized using remote sensing techniques (You *et al*., 2017) [123]. This facilitates identification of crops grown, detecting land-cover change, and the assessment of the level of expansion in agriculture versus degradation. Information helps policymakers and farmers better manage agricultural landscapes to maximize and optimize land use (Joshi *et al*., 2016) [48].
- **Monitoring weather and climate:** Through satellites, long-term trends in temperature, precipitation, and wind regimes, which directly impact agricultural productivity, can be tracked. It enables prediction of climatic anomalies like drought or floods that would lead farmers to make changes in practice (Dewitte, *et al*., 2021) [25].



Source: Pokhariyal *et al*., 2023) [91]

Fig 5: Digital transformation in agriculture representing integration of remotely sensed data, crop simulation, and machine learning

3.2 Key remote sensing datasets

Several satellite and aerial platforms are now used to provide high-resolution data that supports diverse agricultural applications (Zhang *et al*., 2020) [124]. These datasets enable farmers, researchers, and agricultural practitioners to assess crop health, optimize resource use, and forecast yields:

- **Landsat Program:** Landsat satellites are a major part of monitoring agriculture for generations, even offering multispectral and thermal imagery (Wulder *et al*., 2022) [119]. These datasets are critical for long-term analysis in vegetation, mapping land cover, and studying trends in agriculture over time (Wulder *et al*., 2019) [118].

- **Sentinel Satellites (Sentinel-2):** Sentinel-2 offers high-resolution optical and near-infrared imagery, enabling detailed analysis of vegetation health (Drusch *et al*., 2012) [26]. This data is used for crop condition monitoring, identifying plant stress, and assessing yield potential. Its ability to capture images every 5 days ensures frequent updates for dynamic agricultural systems.
- **MODIS:** This sensor installed in NASA's satellites captures low-resolution data good for large area vegetation monitoring, assessment of drought and climate studies (Sheffield *et al*., 2018) [107]. It offers a nearly global coverage that is also quite useful to identify the large-area trends of vegetation (Ndehedehe *et al*., 2022) [80].
- **UAV (Unmanned Aerial Vehicles):** Drones equipped with multispectral and hyperspectral sensors allow for high-resolution images to be interpreted at the field level. These are now common in precision agriculture for crop health monitoring, optimization of irrigation, or identification of early pest or disease outbreaks at farm level (Adão *et al*., 2017) [1].

3.3 Vegetation Indices for Yield Prediction

- **Vegetation indices (VIs):** Derived from RS data represent an essential tool for crop health assessment and yield prediction in agriculture (Karthikeyan *et al*., 2020) [51]. Key indices used for yield prediction include vegetation indices based on light reflectance at different wavelengths:
- **NDVI:** Stands for Normalized Difference Vegetation Index. This is one of the most frequently used VIs in assessing the general health of the plants. NDVI measures the difference between near-infrared and red light reflected from the plants. It is positively correlated with chlorophyll and plant vigor. NDVI can be used to determine crop health, so proper harvesting time is determined.
- **Enhanced Vegetation Index (EVI):** EVI is more accurate than NDVI since it removes the effect of clouds, aerosols, and soil brightness. EVI, therefore, becomes more accurate in densely vegetated areas rather than NDVI because the latter becomes saturated (Matsushita *et al*., 2007) [73]. The use of EVI has become significant in high-density cropped areas to assess the growth and stress of crop vegetation (Gurung *et al*., 2009) [41].

Table 2: Yield prediction using RS and ML techniques in India

Crop	Input Functionality	Algorithm	Scale	Result	Reference
Rice	RS variable: NDVI Satellite: Sentinel 1A and Sentinel 2B	RF (Random Forest)	Regional	Rice yield predicted as 1.60 metric tons/hectare	(Ranjan *et al*., 2019) [95]
Sugarcane	(a) RS variable: NDVI (b) Meteorological: temperature, dew point temperature, soil temperature, soil moisture, precipitation, relative humidity, sunshine duration, evapotranspiration Satellite: LANDSAT 8	SVR, LR, NB, DT, SVR-RBF	Regional	Best model: SVR-RBF, OA = 83.49%	(Medar *et al*., 2019) [74]
Rice	RS variable: NDVI Meteorological data: RF, Temperature, SR, RH Satellite: LANDSAT 8	ANN, RF	Regional	Combined ANN models with Boruta feature selection and random forest importance: 0.651-0.663	(Chandra *et al*., 2019) [18]
Wheat	(a) RS variable: NDVI, NIRv, NDWI (b) Meteorological: Tmin, Tmax, Tmean, RF, VPD, SWdown, day length Satellite: MODIS	CNN, RF, RR	IGP	Best model: CNN with NSE of 0.868	(Wolanin *et al*., 2020) [117]
Sugarcane	RS variable: RVI, NDVI, SAVI, OSAVI, DVI, GCI, EVI, ARVI, VARI, NDMI, NDWI, NN, NG Satellite: LANDSAT-8	SVR, CART, k-NN, RF	Regional	Best model: RF with $R^2 = 0.94$ and RMSE = 1.51 t/ha	(Singla *et al*., 2020) [110]
Sugarcane	RS variable: NDVI, WSI, APAR, CWSI (NDVI/Ts) Satellite: LANDSAT 8	ANN	Field	Estimated sugarcane yield with 95% accuracy	(Krupavathi *et al*., 2021) [61]
Cotton	(a) RS variable: NDVI (b) Meteorological: RF Satellite: MODIS	RF	Regional	September month: $R^2 = 0.69$	(Prasad *et al*., 2021) [93]
Rice	(a) RS variable: MODIS-LAI (b) Observed yields (2003-2017) from the Directorate of Economics and Statistics (MoA) Satellite: MODIS	GBR	Regional	Validation for 2016: $R^2 = 0.84$ Validation for 2017: $R^2 = 0.77$	(Arumugam *et al*., 2021) [10]
Sugarcane	RS variable: NDVI, EVI, LAI, FPAR, ET, PET, LE, GPP, RF Satellite: MODIS, CHIRPS	RF, SVR, GBR, XGB	Regional	Best model: GBR, $R^2 = 0.66$, RMSE = 7.15 t/ha	Nihar *et al*., 2022 [82]

Source: Pokhariyal *et al*., 2023 [91]

3.5 Challenges in Remote Sensing

Despite the numerous advantages, remote sensing in agriculture faces certain challenges.

- **Data Resolution and Availability:** High-resolution RS data is often expensive and not readily available to small-scale farmers. The cost of acquiring and processing high-quality data is a significant barrier to its widespread adoption, especially in developing regions (Kempeneers *et al*., 2011) [54].
- **Cloud Cover and Weather Conditions:** Weather conditions in general, cloud cover, and the lack of clear satellite imagery may prevent or limit the data-gathering process (Alavipanah *et al*., 2010) [5]. Especially where rainfall or cloud cover is common, it becomes quite challenging to get clear images, as farmers need to monitor crops continuously.
- This needs expertise in remote sensing, image processing, and analysis of RS data. Innumerable farmers and practicing

agriculturalists don't possess adequate technical knowledge and facilities to be able to operate the RS tools properly and so, their integration at the field levels is difficult.

4. Applications of ML for Crop Yield Predictions
The integration of machine learning with remote sensing data has the potential to significantly improve the accuracy and efficiency of crop yield prediction.

4.1 Workflow for Integration
- **Data Collection:** High-quality RS data, along with ground-truth data from field observations, form the foundation of ML models for agricultural forecasting (Khilola, 2024) [57].
- **Preprocessing:** The RS data generally needs preprocessing such as noise removal, feature extraction, and data normalization for the suitability of machine learning algorithms (Sishodia *et al*., 2020) [111].
- **Model Development:** Machine learning models are developed using historical and real-time RS data to predict various outcomes, such as crop yield and health (El-Omairi *et al*., 2023) [30].
- **Validation and Deployment:** The performance of ML models is validated using independent data sets, and the models are deployed for real-time predictions.

4.2 Success Stories and Applications
- **Wheat Yield Prediction:** Researchers have used Sentinel-2 imagery with Random Forest models to predict wheat yields with high accuracy. This approach helps in forecasting regional yields and informs policy decisions (Segarra *et al*., 2022) [101].
- **Rice Yield Estimation:** UAV imagery combined with CNNs has been used to estimate rice yield in precision farming systems. This technique enhances yield forecasting at the field level (Yang *et al*., 2019) [121].

5. Challenges and Limitations
5.1 Quality and Availability of Data
The high-resolution RS data is often costly, and a large amount of data requires massive storage and processing capabilities. In addition, data quality depends on the platform, sensor type, and environmental conditions that may affect the reliability of the results.

5.2 Algorithm Complexity
The selection and calibration of the appropriate ML algorithms for specific agricultural applications are difficult. Algorithms have to be fine-tuned for the specific difficulties that different crops, climates, and regions may pose.

5.3 Environmental and Regional Variability
Agricultural models that rely on RS data have to take into consideration the variability in local climates, soil types, and topography. Regional variations can greatly influence the performance of predictive models and thus require calibrations according to local conditions.

6. Future Directions
Future perspectives of RS in agriculture
Advancements in technology are very promising for RS. Several developments lie on the horizon.

- **Transfer learning models:** With the transfer learning model, it would be possible to adapt the data of RS with machine learning algorithms of the RS to regions. This way, the entire set of the developed technology could easily reach the far-off farmers.
- **Integration with IoT:** The integration of RS with Internet of Things (IoT) technologies will enable real time sensing of data collected by sensors in the field, which will help farmers attain real-time information regarding crop health, soil moisture, and environmental conditions.
- **Hyperspectral imaging advancement:** The hyperspectral imaging technology captures a larger range of wavelengths than the multispectral sensors, thereby providing even more detailed insights about crop health, nutrient status, and stress factors. This may have the potential to be even more precise in crop management and yield prediction.

7. Conclusion
The integration of machine learning and remote sensing has transformative potential for agriculture, especially in improving crop yield prediction and precision farming. Challenges remain in terms of data resolution, cost, and technical expertise, but continuous advancement in technology, data processing, and accessibility will overcome these barriers. As these technologies evolve, they promise to drive more sustainable and efficient agricultural practices, ultimately benefiting farmers, researchers, and the global food supply chain.
The integration of ML and RS offers transformative potential in agriculture, particularly for crop yield prediction. While challenges exist, continuous advancements in technology and data availability are expected to bridge the gaps, paving the way for sustainable agricultural practices.

8. References

1. Adão T, Hruška J, Pádua L, Bessa J, Peres E, Morais R, *et al*. Hyperspectral imaging: A review on UAV-based sensors, data processing and applications for agriculture and forestry. Remote Sens. 2017;9(11):1110.
2. Ahmed M, Asim M, Ahmad S, Aslam M. Climate change, agricultural productivity, and food security. In: Global agricultural production: Resilience to climate change. Cham: Springer International Publishing, 2023, p. 31-72.
3. Akintuyi OB. Adaptive AI in precision agriculture: A review: investigating the use of self-learning algorithms in optimizing farm operations based on real-time data. Res J Multidiscip Stud. 2024;7(02):016-030.
4. Alahmad T, Neményi M, Nyéki A. Applying IoT sensors and big data to improve precision crop production: A review. Agronomy. 2023;13(10):2603.
5. Alavipanah SK, Matinfar HR, Emam RA, Khodaei K, Hadji Bagheri R, Panah YA. Criteria of selecting satellite data for studying land resources. Desert. 2010;15(2):83-102.
6. Alzubi J, Nayyar A, Kumar A. Machine learning from theory to algorithms: An overview. J Phys: Conf Ser. 2018;1142:012012.
7. Amiri Z, Heidari A, Navimipour NJ. Comprehensive survey of artificial intelligence techniques and strategies for climate change mitigation. Energy. 2024;132827.
8. Argento F, Anken T, Abt F, Vogelsanger E, Walter A, Liebisch F. Site-specific nitrogen management in winter wheat supported by low-altitude remote sensing and soil data. Precis Agric. 2021;22:364-386.
9. Aria M, Cuccurullo C, Gnasso A. A comparison among interpretative proposals for Random Forests. Mach Learn Appl. 2021;6:100094.
10. Arumugam P, Chemura A, Schauberger B, Gornott C.

Remote sensing-based yield estimation of rice (*Oryza sativa* L.) using gradient boosted regression in India. Remote Sens. 2021;13(12):2379. Available from: https://doi.org/10.3390/rs13122379
11. Badapanda K, Mishra DP, Salkuti SR. Agriculture data visualization and analysis using data mining techniques: application of unsupervised machine learning. TELKOMNIKA (Telecommunication Comput Electron Control). 2022;20(1):98-108.
12. Badshah A, Alkazemi BY, Din F, Zamli KZ, Haris M. Crop classification and yield prediction using robust machine learning models for agricultural sustainability. IEEE Access, 2024.
13. Bansod B, Singh R, Thakur R, Singhal G. A comparison between satellite-based and drone-based remote sensing technology to achieve sustainable development: A review. J Agric Environ Int Dev (JAEID). 2017;111(2):383-407.
14. Bhanumathi S, Vineeth M, Rohit N. Crop yield prediction and efficient use of fertilizers. IEEE Xplore. 2019. Available from: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8698087
15. Bhat SA, Huang NF. Big data and AI revolution in precision agriculture: Survey and challenges. IEEE Access. 2021;9:110209-110222.
16. Bhojani SH, Bhatt N. Wheat crop yield prediction using new activation functions in neural networks. Neural Comput Appl. 2020;32(17):13941-13951. Available from: https://doi.org/10.1007/s00521-020-04797-8
17. Bouguettaya A, Zarzour H, Kechida A, Taberkit AM. Deep learning techniques to classify agricultural crops through UAV imagery: A review. Neural Comput Appl. 2022;34(12):9511-9536.
18. Chandra A, Mitra P, Dubey SK, Ray SS. Machine learning approach for kharif rice yield prediction integrating multi-temporal vegetation indices and weather and non-weather variables. Int Arch Photogrammetry Remote Sens Spatial Inf Sci. 2019;42:187-194.
19. Charbuty B, Abdulazeez A. Classification based on decision tree algorithm for machine learning. J Appl Sci Technol Trends. 2021;2(01):20-28.
20. Charoen-Ung P, Mittrapiyanuruk P. Sugarcane yield grade prediction using random forest and gradient boosting tree techniques. In: 2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE), 2018, p. 1-5. Available from: https://doi.org/10.1109/jcsse.2018.8457391
21. Chlingaryan A, Sukkarieh S, Whelan B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Comput Electron Agric. 2018;151:61-69.
22. Choudhary NK, Chukkapalli SSLL, Mittal S, Gupta M, Abdelsalam M, Joshi A. YieldPredict: A crop yield prediction framework for smart farms. IEEE Xplore, 2020. Available from: https://ieeexplore.ieee.org/abstract/document/9377832
23. Conradt T. Choosing multiple linear regressions for weather-based crop yield prediction with ABSOLUT v1. 2 applied to the districts of Germany. Int J Biometeorol. 2022;66(11):2287-2300.
24. Cravero A, Pardo S, Galeas P, Fenner LJ, Caniupán M. Data type and data sources for agricultural big data and machine learning. Sustainability. 2022;14(23):16131.
25. Dewitte S, Cornelis JP, Müller R, Munteanu A. Artificial intelligence revolutionises weather forecast, climate monitoring and decadal prediction. Remote Sens. 2021;13(16):3209.
26. Drusch M, Del Bello U, Carlier S, Colin O, Fernandez V, Gascon F, *et al*. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sens Environ. 2012;120:25-36.
27. Eddamiri S, Bouras EH, Amazirh A, Hakam O, Ayugi BO, Ongoma V. Modeling the impact of climate change on wheat yield in Morocco based on stacked ensemble learning. Model Earth Syst Environ. 2024;10(5):6413-6433.
28. Elavarasan D, Vincent DR. Reinforced XGBoost machine learning model for sustainable intelligent agrarian applications. J Intell Fuzzy Syst. 2020;39(5):7605-7620.
29. Elbasi E, Zaki C, Topcu AE, Abdelbaki W, Zreikat AI, Cina E, *et al*. Crop prediction model using machine learning algorithms. Appl Sci. 2023;13(16):9288.
30. El-Omairi MA, El Garouani A. A review on advancements in lithological mapping utilizing machine learning algorithms and remote sensing data. Heliyon, 2023.
31. Fahrmeir L, Kneib T, Lang S, Marx B. Regression models. In: Springer Berlin Heidelberg, 2013, p. 21-72.
32. Gandhi N, Armstrong LJ, Petkar O, Tripathy AK. Rice crop yield prediction in India using support vector machines. In: 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), 2016, p. 1-5.
33. Ge Y, Thomasson JA, Sui R. Remote sensing of soil properties in precision agriculture: A review. Front Earth Sci. 2011;5:229-238.
34. Giedion S. Space, time and architecture: the growth of a new tradition. Cambridge, MA: Harvard University Press, 2009.
37. Golubovic N, Krintz C, Wolski R, Sethuramasamyraja B, Liu B. A scalable system for executing and scoring K-means clustering techniques and its impact on applications in agriculture. Int J Big Data Intell. 2019;6(3-4):163-175.
38. Greenacre M, Groenen PJ, Hastie T, d'Enza AI, Markos A, Tuzhilina E. Principal component analysis. Nat Rev Methods Primers. 2022;2(1):100.
39. Gupta G, Setia R, Meena A, Jaint B. Environment monitoring system for agricultural application using IoT and predicting crop yield using various data mining techniques. IEEE Xplore. Available from: https://ieeexplore.ieee.org/abstract/document/9138032
40. Gupta M, Krishna BVS, Kavyashree B, Narapureddy HR, Surapaneni N, Varma K. Crop yield prediction techniques using machine learning algorithms. 2022 8th International Conference on Smart Structures and Systems (ICSSS), 2022. Available from: https://doi.org/10.1109/icsss54381.2022.9782246
41. Gurung RB, Breidt FJ, Dutin A, Ogle SM. Predicting Enhanced Vegetation Index (EVI) curves for ecosystem modeling applications. Remote Sens Environ. 2009;113(10):2186-2193.
42. Han W, Zhang X, Wang Y, Wang L, Huang X, Li J, *et al*. A survey of machine learning and deep learning in remote sensing of geological environment: Challenges, advances, and opportunities. ISPRS J Photogramm Remote Sens. 2023;202:87-113.
43. Houborg R, Fisher JB, Skidmore AK. Advances in remote sensing of vegetation function and traits. Int J Appl Earth Obs Geoinf. 2015;43:1-6.
44. Imade A, Ounacer S, Ghoumari ELMY, Ardchir S, Azzouazi M. Agricultural yield prediction using ML

algorithms in the Industry 5.0. In: Industry 5.0 and Emerging Technologies: Transformation Through Technology and Innovations. Cham: Springer Nature Switzerland, 2024, p. 135-157.
45. Iqbal N, Shahzad MU, Sherif ESM, Tariq MU, Rashid J, Le TV, *et al*. Analysis of wheat-yield prediction using machine learning models under climate change scenarios. Sustainability. 2024;16(16):6976.
46. Jewiss JL, Brown ME, Escobar VM. Satellite remote sensing data for decision support in emerging agricultural economies: How satellite data can transform agricultural decision making [Perspectives]. IEEE Geosci Remote Sens Mag. 2020;8(4):117-133.
47. Jin X, Kumar L, Li Z, Feng H, Xu X, Yang G. Remote sensing-based biomass estimation and its spatio-temporal variations in grassland: A review. Ecol Indic. 2018;60:123-131.
48. Joshi N, Baumann M, Ehammer A, Fensholt R, Grogan K, Hostert P, *et al*. A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. Remote Sens. 2016;8(1):70.
49. Kamath P, Patil P, Sushma S. Crop yield forecasting using data mining. Glob Transitions Proc, 2021. Available from: https://doi.org/10.1016/j.gltp.2021.08.008
50. Kamilaris A, Prenafeta-Boldú FX. Deep learning in agriculture: A survey. Comput Electron Agric. 2018;147:70-90.
51. Karthikeyan L, Chawla I, Mishra AK. A review of remote sensing applications in agriculture for food security: Crop growth and yield, irrigation, and crop losses. J Hydrol. 2020;586:124905.
52. Karunathilake EMBM, Le AT, Heo S, Chung YS, Mansoor S. The path to smart farming: Innovations and opportunities in precision agriculture. Agriculture. 2023;13(8):1593.
53. Keerthana M, Meghana KJM, Pravallika S, Kavitha M. An ensemble algorithm for crop yield prediction. 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021. Available from: https://doi.org/10.1109/icicv50876.2021.9388479
54. Kempeneers P, Sedano F, Seebach L, Strobl P, San-Miguel-Ayanz J. Data fusion of different spatial resolution remote sensing images applied to forest-type mapping. IEEE Trans Geosci Remote Sens. 2011;49(12):4977-4986.
55. Khaki S, Pham H, Wang L. Simultaneous corn and soybean yield prediction from remote sensing data using deep transfer learning. Sci Rep. 2021;11(1):11132. Available from: https://doi.org/10.1038/s41598-021-89779-z
56. Khaki S, Wang L, Archontoulis SV. A CNN-RNN framework for crop yield prediction. Front Plant Sci. 2020;10:1750.
57. Khilola A. Crop yield prediction using machine learning, multi-source remote sensing technologies, and data fusion: A case study of Mezőhegyes Hungary. [Doctoral dissertation]. Szeged University (Hungary), 2024.
58. Kim N, Ha KJ, Park NW, Cho J, Hong S, Lee YW. A comparison between major artificial intelligence models for crop yield prediction: Case study of the Midwestern United States, 2006-2015. ISPRS Int J Geo-Inf. 2019;8(5):240. Available from: https://doi.org/10.3390/ijgi8050240
59. Krasnopolsky VM, Rabinovitz FMS. Complex hybrid models combining deterministic and machine learning components for numerical climate modeling and weather prediction. Neural Netw. 2006;19(2):122-134.
60. Krupavath K, Babu MR, Mani A. Comparative evaluation of neural networks in crop yield prediction of paddy and sugarcane crop. In: The Digital Agricultural Revolution. Wiley, 2022, p. 25-55. Available from: https://doi.org/10.1002/9781119823469.ch2
61. Krupavathi K, Raghubabu M, Mani A, Prasad PRK, Edukondalu L. Field-scale estimation and comparison of the sugarcane yield from remote sensing data: A machine learning approach. J Indian Soc Remote Sens. 2021;50(2):299-312. Available from: https://doi.org/10.1007/s12524-021-01355-6
62. Kulyal M, Saxena P. Machine learning approaches for crop yield prediction: A review. In: 2022 7th International Conference on Computing, Communication and Security (ICCCS), 2022, p. 1-7. IEEE.
63. Kumar YJN, Spandana V, Vaishnavi VS, Neha K, Devi VGRR. Supervised machine learning approach for crop yield prediction in agriculture sector. 2020 5th International Conference on Communication and Electronics Systems (ICCES), 2020. Available from: https://doi.org/10.1109/icces48766.2020.9137868
64. Lawley V, Lewis M, Clarke K, Ostendorf B. Site-based and remote sensing methods for monitoring indicators of vegetation condition: An Australian review. Ecol Indic. 2016;60:1273-1283.
65. Leo S, Migliorati DAM, Grace PR. Predicting within-field cotton yields using publicly available datasets and machine learning. Agron J. 2021;113(2):1150-1163.
66. Li H, Mei X, Wang J, Huang F, Hao W, Li B. Drip fertigation significantly increased crop yield, water productivity and nitrogen use efficiency with respect to traditional irrigation and fertilization practices: A meta-analysis in China. Agric Water Manag. 2021;244:106534.
67. Li L, Liu L, Peng Y, Su Y, Hu Y, Zou R. Integration of multimodal data for large-scale rapid agricultural land evaluation using machine learning and deep learning approaches. Geoderma. 2023;439:116696.
68. Liaghat S, Balasundram SK. A review: The role of remote sensing in precision agriculture. Am J Agric Biol Sci. 2010;5(1):50-55.
69. Lobell DB, Thau D, Seifert C, Engle E, Little B. A scalable satellite-based crop yield mapper. Remote Sens Environ. 2015;164:324-333.
70. Manivasagam MA, Sumalatha P, Likitha A, Pravallika V, Satish KV, Sreeram S. An efficient crop yield prediction using machine learning. Int J Res Eng Sci Manag. 2022;5(3):106-111. Available from: https://journals.resaim.com/ijresm/article/view/1862
73. Matsushita B, Yang W, Chen J, Onda Y, Qiu G. Sensitivity of the enhanced vegetation index (EVI) and normalized difference vegetation index (NDVI) to topographic effects: a case study in high-density cypress forest. Sensors. 2007;7(11):2636-2651.
74. Medar RA, Rajpurohit VS, Ambekar AM. Sugarcane crop yield forecasting model using supervised machine learning. Int J Intell Syst Appl. 2019;11(8):11.
75. Mehta P, Shah H, Kori V, Vikani V, Shukla S, Shenoy M. Survey of unsupervised machine learning algorithms on precision agricultural data. In: 2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), IEEE, 2015 Mar, p. 1-8.
76. Tovar MCL. NDVI as indicator of degradation. Unasylva. 2011;62(238):39-46.

77. Mohamed M. Agricultural sustainability in the age of deep learning: current trends, challenges and future Trajectories. Sustainable Machine Intelligence J. 2023;4:2-1.
78. Mohyuddin G, Khan MA, Haseeb A, Mahpara S, Waseem M, Saleh AM. Evaluation of Machine Learning approaches for precision farming in Smart Agriculture System-A comprehensive Review. IEEE Access, 2024.
79. Moskolaï WR, Abdou W, Dipanda A, Kolyang. Application of deep learning architectures for satellite image time series prediction: A review. Remote Sens. 2021;13(23):4822.
80. Ndehedehe C. Satellite remote sensing of terrestrial hydrology. Springer, 2022.
81. Ngie A, Ahmed F. Estimation of maize grain yield using multispectral satellite data sets (SPOT 5) and the random forest algorithm. S Afr J Geomatics. 2018;7(1):11-30.
82. Nihar A, Patel NR, Danodia A. Machine-learning-based regional yield forecasting for sugarcane crop in Uttar Pradesh, India. J Indian Soc Remote Sens. 2022;50(6):1519-1530. DOI: 10.1007/s12524-022-01553-2.
83. Nishant PS, Sai Venkat P, Avinash BL, Jabber B. Crop yield prediction based on Indian agriculture using machine learning. IEEE Xplore, 2020. Available from: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9154036
84. Nzekwe CJ, Kim S, Mostafa SA. Interaction selection and prediction performance in high-dimensional data: A comparative study of statistical and tree-based methods. J Data Sci. 2024;22(2).
85. Ozdogan M, Yang Y, Allez G, Cervantes C. Remote sensing of irrigated agriculture: Opportunities and challenges. Remote Sens. 2010;2(9):2274-2304.
86. Pantazi XE, Moshou D, Alexandridis T, Whetton RL, Mouazen AM. Wheat yield prediction using machine learning and advanced sensing techniques. Comput Electron Agric. 2016;121:57-65.
87. Pascucci S, Carfora MF, Palombo A, Pignatti S, Casa R, Pepe M, *et al*. A comparison between standard and functional clustering methodologies: Application to agricultural fields for yield pattern assessment. Remote Sens. 2018;10(4):585.
88. Patil D. Artificial Intelligence Innovations in Precision Farming: Enhancing Climate-Resilient Crop Management. SSRN, 2024. Available from: https://ssrn.com/abstract=5057424
89. Paudel D, Boogaard H, Wit DA, Janssen S, Osinga S, Pylianidis C, *et al*. Machine learning for large-scale crop yield forecasting. Agric Syst. 2021;187:103016.
90. Peng M, Liu Y, Khan A, Ahmed B, Sarker SK, Ghadi YY, *et al*. Crop monitoring using remote sensing land use and land change data: Comparative analysis of deep learning methods using pre-trained CNN models. Big Data Res. 2024;36:100448.
91. Pokhariyal S, Patel NR, Govind A. Machine learning-driven remote sensing applications for agriculture in India: A systematic review. Agronomy. 2023;13(9):2302.
92. Polwaththa KPGDM, Amarasinghe STC, Amarasinghe AAYD, Amarasinghe AAY. Exploring Artificial Intelligence and Machine Learning in Precision Agriculture: A Pathway to Improved Efficiency and Economic Outcomes in Crop Production. Am J Agric Sci Eng Technol. 2024;8(3):50-59.
93. Prasad NR, Patel NR, Danodia A. Crop yield prediction in cotton for regional level using random forest approach. Spatial Inf Res. 2021;29(2):195-206. DOI: 10.1007/s41324-020-00344-7.
94. Priya R, Ramesh D, Khosla E. Crop prediction on the region belts of India: A naïve Bayes MapReduce precision agricultural model. In: 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2018, p. 355-360. DOI: 10.1109/icacci.2018.8554948.
95. Ranjan AK, Parida BR. Paddy acreage mapping and yield prediction using sentinel-based optical and SAR data in Sahibganj district, Jharkhand (India). Spatial Inf Res. 2019;27(4):399-410.
96. Rashid M, Bari BS, Yusup Y, Kamaruddin MA, Khan N. A comprehensive review of crop yield prediction using machine learning approaches with special emphasis on palm oil yield prediction. IEEE Access. 2021;9:63406-63439.
97. Sabir RM, Mehmood K, Sarwar A, Safdar M, Muhammad NE, Gul N, *et al*. Remote sensing and precision agriculture: A sustainable future. In: Transforming Agricultural Management for a Sustainable Future: Climate Change and Machine Learning Perspectives. Cham: Springer Nature Switzerland, 2024, p. 75-103.
98. Saiz-Rubio V, Más RF. From smart farming towards agriculture 5.0: A review on crop data management. Agronomy. 2020;10(2):207.
99. Sakamoto T. Incorporating environmental variables into a MODIS-based crop yield estimation method for United States corn and soybeans through the use of a random forest regression algorithm. ISPRS J Photogramm Remote Sens. 2020;160:208-228.
100. Santos CFGD, Papa JP. Avoiding overfitting: A survey on regularization methods for convolutional neural networks. ACM Comput Surv. 2022;54(10s):1-25.
101. Segarra J, Araus JL, Kefauver SC. Farming and Earth Observation: Sentinel-2 data to estimate within-field wheat grain yield. Int J Appl Earth Obs Geoinf. 2022;107:102697.
102. Segarra J, Buchaillot ML, Araus JL, Kefauver SC. Remote sensing for precision agriculture: Sentinel-2 improved features and applications. Agronomy. 2020;10(5):641.
103. Shah A, Dubey A, Hemnani V, Gala D, Kalbande DR. Smart farming system: Crop yield prediction using regression techniques. In: Lecture Notes on Data Engineering and Communications Technologies, 2018, p. 49-56. DOI: 10.1007/978-981-10-8339-6_6.
104. Shahhosseini M, Hu G, Huber I, Archontoulis SV. Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. Sci Rep. 2021;11(1):1606. DOI: 10.1038/s41598-020-80820-1.
105. Sharma K, Shivandu SK. Integrating artificial intelligence and internet of things (IoT) for enhanced crop monitoring and management in precision agriculture. Sensors Int. 2024;100292.
106. Sharma R, Kamble SS, Gunasekaran A, Kumar V, Kumar A. A systematic literature review on machine learning applications for sustainable agriculture supply chain performance. Comput Oper Res. 2020;119:104926.
107. Sheffield J, Wood EF, Pan M, Beck H, Coccia G, Capdevila SA, *et al*. Satellite remote sensing for water resources management: Potential for supporting sustainable development in data-poor regions. Water Resour Res. 2018;54(12):9724-9758.
108. Singh A, Khemka P, Prasad R. Advances in remote sensing for crop monitoring and yield prediction: A comprehensive review. Agric Syst. 2020;179:102784.
109. Singh RK, Berkvens R, Weyn M. AgriFusion: An

architecture for IoT and emerging technologies based on a precision agriculture survey. IEEE Access. 2021;9:136253-136283.
110. Singla SK, Garg RD, Dubey OP. Ensemble machine learning methods to estimate the sugarcane yield based on remote sensing information. Rev Intell Artif. 2020;34(6):731-743. DOI: 10.18280/ria.340605.
111. Sishodia RP, Ray RL, Singh SK. Applications of remote sensing in precision agriculture: A review. Remote Sens. 2020;12(19):3136.
112. Titirmare S, Margal PB, Gupta S, Kumar D. AI-powered predictive analytics for crop yield optimization. In: Agriculture 4.0. CRC Press, 2024, p. 89-110.
113. Tiwari P, Shukla P. Artificial neural network-based crop yield prediction using NDVI, SPI, VCI feature vectors. In: Information and Communication Technology for Sustainable Development. Springer, 2019, p. 585-594. DOI: 10.1007/978-981-13-7166-0_58.
114. Wang J, Wang P, Tian H, Tansey K, Liu J, Quan W. A deep learning framework combining CNN and GRU for improving wheat yield estimates using time series remotely sensed multi-variables. Comput Electron Agric. 2023;206:107705.
115. Wang J, Wang Y, Li G, Qi Z. Integration of remote sensing and machine learning for precision agriculture: A comprehensive perspective on applications. Agronomy. 2024;14(9):1975.
116. Weiss M, Jacob F, Duveiller G. Remote sensing for agricultural applications: A meta-review. Remote Sens Environ. 2020;236:111402.
117. Wolanin A, García MG, Valls CG, Chova GL, Meroni M, Duveiller G, *et al*. Estimating and understanding crop yields with explainable deep learning in the Indian Wheat Belt. Environ Res Lett. 2020;15(2):024019. DOI: 10.1088/1748-9326/ab5d3e.
118. Wulder MA, Loveland TR, Roy DP, Crawford CJ, Masek JG, Woodcock CE, *et al*. Current status of Landsat program, science, and applications. Remote Sens Environ. 2019;225:127-147.
119. Wulder MA, Roy DP, Radeloff VC, Loveland TR, Anderson MC, Johnson DM, *et al*. Fifty years of Landsat science and impacts. Remote Sens Environ. 2022;280:113195.
120. Yadav N, Sidana N. Precision Agriculture Technologies: Analysing the use of advanced technologies, such as drones, sensors, and gps, in precision agriculture for optimizing resource management, crop monitoring and yield prediction. J Adv Zool, 2023, p. 44.
121. Yang Q, Shi L, Han J, Zha Y, Zhu P. Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images. Field Crops Res. 2019;235:142-153.
122. Yang W, Kogan F, Guo W. An ongoing blended long-term vegetation health product for monitoring global food security. Agronomy. 2020;10(12):1936.
123. You J, Li X, Low M, Lobell D, Ermon S. Deep Gaussian process for crop yield prediction based on remote sensing data. Proceedings of the AAAI Conference on Artificial Intelligence. 2017;31(1):4559-4566.
124. Zhang C, Marzougui A, Sankaran S. High-resolution satellite imagery applications in crop phenotyping: An overview. Comput Electron Agric. 2020;175:105584.
125. Zhou W, Yan Z, Zhang L. A comparative study of 11 non-linear regression models highlighting autoencoder, DBN, and SVR, enhanced by SHAP importance analysis in soybean branching prediction. Sci Rep. 2024;14(1):5905.
126. Zhou ZH. Ensemble methods: foundations and algorithms. CRC Press, 2025.