STM JOURNALS

Research & Reviews: Journal of
Agricultural Science and Technology

https://journals.stmjournals.com/rrjoast

ISSN: 2278-2206 (Online)
ISSN: 2349-3682 (Print)
Volume 14, Issue 3, 2025
September–December
DOI (Journal): 10.37591/RRJoAST

Review RRJoAST

Advancements in Agricultural Forecasting: A Review of Machine Learning Based Crop Yield Prediction

Vijay Laxmi[1], Gaurav Singla[2,*]

Abstract

Agricultural productivity plays a critical role in global food security. Accurate prediction of crop yield is essential for optimizing resource allocation and decision-making in farming. The rapid advancements in machine learning and deep learning have transformed agricultural forecasting, enabling data-driven approaches for crop prediction. The present review paper provided a comprehensive analysis of various machine learning and deep learning techniques applied in crop yield forecast, highlighting the ineffectiveness, challenges, and future directions. The study explored different models, including Random Forest, Support Vector Technologies, Artificial Neural Networks, Long Short-Term Memory Networks, and other ensemble methods, comparing the performance based on accuracy metrics such as R^2 score, root mean squared error, and mean absolute error. Additionally, the present paper discussed the incorporation of Internet of Things and remote sensing technologies in modern precision agriculture. Statistical methods, while deep learning models excel in handling complex, nonlinear relationships in agricultural data. However, challenges such as data availability, environmental variability, and computational efficiency remain key barriers. The present review aimed to provide insights into the potential of artificial intelligence-driven approaches in enhancing agricultural sustainability and precision farming, paving the way for future research and innovations in smart agriculture.

Keywords: Crop yield forecasting, machine learning in agriculture, deep learning for crop prediction, precision agriculture, smart farming, artificial intelligence in agriculture

INTRODUCTION

Agriculture is a fundamental sector that sustains human civilization [1], playing a vital role in food production, economic development, and environmental sustainability. With the rapid growth of the worldwide population, estimated to reach 9.7 billion by the year 2050, the demand for agricultural crops is increasing at an unprecedented rate [2]. However, this rising demand is met with significant trials, including climate change, soil degradation, water scarcity, and random weather patterns, all of which directly impact crop yields [3]. Outdated methods of crop yield prediction, which primarily rely on farmer's experience, historical trends, and basic statistical models, often fail to provide accurate and timely [1] insights due to the complexity of modern agricultural systems [4]. As a result, researchers and agricultural experts are turning to advanced computational techniques, particularly Machine Learning (ML) and Deep Learning (DL), to enhance the correctness and competence of crop yield forecasting [5]. ML and artificial intelligence (AI) have emerged as transformative tools in precision agriculture, offering the skill to analyze vast datasets, identify hidden patterns, and generate predictive models for crop production [6]. Various ML algorithms, for example Random Forest (RF), Decision Trees, Support Vector Machines (SVM), and Artificial Neural Networks (ANN), have been

***Author for Correspondence**
Gaurav Singla
E-mail: gauravsingla444@gmail.com

[1]Professor, Faculty of Computing & Guru Kashi University, Talwandi Sabo, Punjab, India
[2]Research Scholar, Faculty of Computing & Guru Kashi University, Talwandi Sabo, Punjab, India

Received Date: July 10, 2025
Accepted Date: July 30, 2025
Published Date: August 20, 2025

Citation: Vijay Laxmi, Gaurav Singla. Advancements in Agricultural Forecasting: A Review of Machine Learning Based Crop Yield Prediction. Research & Reviews: Journal of Agricultural Science and Technology. 2025; 14(3): 32–38p.

employed to estimate yield based on a crowd of factors, including soil composition, rainfall, temperature, and fertilizer usage [2, 4]. Furthermore, DL methods, particularly Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, have demonstrated superior capabilities in handling high-dimensional agricultural data, improving prediction accuracy in complex scenarios [3]. These AI-driven approaches not only aid farmers in optimizing resource allocation but also support policymakers in making informed decisions regarding food security and sustainability [7].

Recent advancements in agricultural forecasting have also been bolstered by the incorporation of remote sensing technologies, the Internet of Things (IoT), and satellite-based data collection [5]. IoT-enabled smart farming techniques utilize real-time sensor data to monitor soil health, crop growth, and weather conditions, allowing for proactive decision-making [6]. Additionally, satellite imagery and drone-based surveillance provide high-resolution agricultural insights, further refining predictive models and enabling precision farming strategies [7]. These technological advancement shave significantly improved the efficiency of farming practices, reducing crop losses, and enhancing productivity [3].

Despite these breakthroughs, several challenges persist in the widespread adoption of AI-driven crop yield prediction models. Data quality and availability remain key concerns, as accurate predictions require extensive datasets covering multiple seasons and diverse geographical regions [4]. Environmental uncertainties, such as unexpected droughts or pest infestations, introduce variability that is difficult to model with traditional ML approaches [5]. Additionally, the computational complexity and high costs associated with training deep learning models pose barriers to their implementation, particularly for small-scale farmers with limited technological access [2]. Addressing these trials requires a multidisciplinary approach, combining AI innovations with agronomic expertise, improved data collection methodologies, and policy support for technology adoption in agriculture [6].

The present review paper aimed to provide a complete analysis of the latest advancements in ML-based crop yield prediction, exploring the effectiveness of various algorithms and methodologies. It examines the role of big data analytics, IoT integration, and remote sensing technologies in enhancing predictive accuracy. By synthesizing insights from recent research, the present paper highlighted key trends, research gaps, and future directions in AI-driven agricultural forecasting. The ultimate goal was to tie the gap between cutting-edge technology and practical farming applications, contributing to the advancement of precision agriculture and sustainable food production systems.

METHODOLOGY

The present review paper followed a systematic approach to analyze recent developments in ML-based crop yield prediction. The methodology involved four key stages—literature collection, categorization of ML techniques, comparative analysis of models, and identification of challenges and future directions. This structured approach ensured a comprehensive [1] understanding of the state-of-the-art technologies in agricultural forecasting.

Literature Collection and Selection

A thorough literature review was conducted by investigating peer-reviewed journal articles, conference papers, and scientific reports published between 2022 and 2024. The primary sources of information were retrieved from reputed files such as Elsevier, IEE explore, MDPI, and Science Direct. The selection criteria for including papers in the present review were:

- Relevance to crop yield forecast using ML and DL techniques.
- Studies presenting comparative analyses of different ML/DL models for crop forecasting.
- Research integrating IoT, remote sensing, and satellite data in crop yield prediction.
- Papers discussing challenges, limitations, and future opportunities in AI-driven agricultural forecasting.

From the collected literature, six major papers were identified as key references for the present review [2–7].

Categorization of Machine Learning Techniques

To systematically evaluate advancements in crop yield prediction [1], ML techniques were categorized into:

Traditional Machine Learning Representations

- Random Forest (RF) [2, 4]
- Support Vector Machines (SVM) [3]
- Decision Trees (DT) [5]
- k-Nearest Neighbors (KNN) [4]

Deep Learning Models

- Artificial Neural Networks (ANN) [6]
- Long Short-Term Memory (LSTM) networks [2]
- Convolutional Neural Networks (CNN) [5]

Hybrid and Ensemble Learning Approaches

- Extreme Gradient Boosting (XGBoost) [4]
- Extra Trees Regressor (ETR) [7]
- Stacked Machine Learning Models [6]

Each of these models was assessed based on accuracy metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R^2 score, which were used for performance assessment in the selected papers.

Comparative Analysis of Models

A comparative evaluation was conducted to analyze the presentation of different ML and DL models in crop yield prediction. The selected studies reported varying accuracy levels based on datasets, environmental factors, and computational techniques:

- RF and Extra Trees Regressor demonstrated the highest prediction accuracy, with R^2 values exceeding 0.96 in most cases [2, 7].
- Deep Learning models (LSTM, ANN, CNN) showed superior performance in handling high-dimensional and time-series data, particularly when integrating climatic and soil variables [5, 6].
- Hybrid models (XGBoost, Stacked Learning) achieved better generalization than stand-alone models, combining the strengths of decision trees and boosting techniques for improved precision [4].

By comparing these models, the present review highlighted the strengths and drawbacks of different approaches, providing insights into their suitability for various agricultural settings.

LITERATURE REVIEW

The increasing demand [1] for food production and the challenges posed by climate change have led to significant advancements in ML and DL for crop yield forecast. Traditional statistical models used for agricultural forecasting often fail to capture the complex, nonlinear relationships among soil properties, weather conditions, and farming practices. In contrast, ML and DL models provide data-driven, adaptive solutions that significantly improve prediction accuracy and decision-making in precision agriculture [2–7].

Machine Learning Models for Crop Yield Prediction

Several studies have explored the use of ML techniques for agricultural forecasting. Jhajharia et al. [2] implemented ML models, including RF, SVM, and Gradient Descent, to forecast crop yields in

Rajasthan, India. Their findings indicated that RF did the best with an R^2 score of 0.963, while other models showed slightly lower accuracy. Similarly, Maheswari and Ramasamy [4] conducted a comparative study on ML techniques and reported that RF and XGBoost consistently out-performed other models such as KKNN and Decision Trees DT.

Deep Learning Methods in Agricultural Forecasting

DL method often failed to account for the complexity of modern agricultural systems. Climate variability, changing soil conditions, and unpredictable market demands require more advanced forecasting models. DL offers the ability to integrate multiple sources of data such as satellite imagery, weather sensors, soil records, and market data, enabling more accurate and timely predictions.

Deep Learning Approaches Used

Several DL architectures have been successfully applied in agricultural forecasting:

- *Convolutional Neural Networks (CNNs):* CNNs are highly effective in analyzing image data. In agriculture, CNNs are applied to satellite and drone images for crop classification, yield prediction, and disease detection. For example, CNN-based models can monitor plant growth and forecast yield by identifying variations in leaf color, plant density, and canopy cover.
- *Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM):* Time-series data such as rainfall patterns, temperature changes, and soil moisture levels are critical for forecasting. RNNs and LSTMs are well-suited for processing sequential data, making them useful in predicting weather conditions and seasonal crop yields. LSTM networks are particularly effective in capturing long-term dependencies, which helps improve the reliability of agricultural forecasts.
- *Generative Adversarial Networks (GANs):* GANs are used to generate synthetic agricultural data where real datasets are limited. For instance, they can simulate weather conditions or soil variations, which can then be used to train predictive models. This helps address the issue of insufficient training data—a common challenge in agricultural applications.
- *Hybrid Models:* Combining DL with traditional ML and statistical models has shown promising results. For example, integrating LSTM with autoregressive models can improve the accuracy of climate-based crop forecasting. Hybrid models can take advantage of the strengths of different techniques while minimizing their weaknesses.

Applications of Deep Learning in Agricultural Forecasting

- *Crop Yield Prediction:* Accurate yield forecasting allows governments and farmers to plan for food storage, distribution, and market pricing.
- *Pest and Disease Forecasting:* DL models can analyze patterns of pest outbreaks and predict future risks, enabling timely preventive measures.
- *Weather and Climate Prediction:* Weather significantly impacts agriculture. DL models enhance the accuracy of rainfall, temperature, and drought predictions.
- *Market Demand Forecasting:* Beyond crop growth, DL techniques can predict consumer demand, helping farmers adjust production accordingly.

Challenges and Limitations

Despite its advantages, the application of DL in agricultural forecasting faces challenges:

- *Data Availability:* Agricultural datasets are often limited, fragmented, or inconsistent.
- *Computational Cost:* Training deep networks requires high computational resources, which may not be accessible in developing regions.
- *Interpretability:* DL models are often seen as "black boxes," making it difficult for farmers to understand the basis of predictions.
- *Infrastructure Gaps:* Poor internet connectivity and lack of advanced equipment in rural areas limit widespread adoption.

Future Prospects

- As data collection technologies improve, DL applications in agriculture are expected to grow significantly. The integration of IoT sensors, drones, and satellite-based monitoring will provide richer datasets for training models. Additionally, efforts are being made to design lightweight and explainable DL models that farmers can use directly through mobile applications. Collaborations between AI researchers, agronomists, and policymakers will further enhance the effectiveness and accessibility of agricultural forecasting systems.

DL models have also been explored for their ability to process large datasets and extract complex patterns. Kolipaka and Namburu [3] implemented LSTM, CNN, and Bi-GRU networks, achieving higher accuracy in time-series crop yield prediction as compared to traditional ML models. Their results highlighted that CNN and Bi-GRU were particularly effective in detecting spatial and temporal dependencies in crop growth data. Additionally, Nikhil et al. [7] analyzed DL models and found that Extra Trees Regressor and LGBM Regressor provided high prediction accuracy, with an R^2 score of 0.9615.

Integration of IoT, Remote Sensing, and Smart Farming

Recent studies emphasized the incorporation of IoT and remote sensing with ML models for real-time crop monitoring. Elbasi et al. [5] developed an IoT- based crop prediction model that used sensor data, satellite imagery, and weather forecasting to improve predictive accuracy. Their study demonstrated that incorporating real-time environmental factors led to significant improvements in crop yield forecasting. Additionally, Patil et al. [6] explored how sensor- driven smart farming techniques could enhance crop selection and resource management.

COMPARISON AND ANALYSIS

Performance Evaluation of ML and DL Models

The studies reviewed demonstrated that RF and ensemble learning methods consistently outperform traditional ML models due to their capability to handle large datasets and complex agricultural variables. Additionally, DL models (CNN, LSTM, Bi-GRU) excel in analyzing time-series and three-dimensional data. The comparison of model performance is summarized in Table 1.

Strengths and Limitations of Approaches

- *RF and XG Boost:* High prediction accuracy and robustness but may require longer computation times for large datasets.
- *Deep Learning (CNN, LSTM, Bi-GRU):* Excellent at capturing spatial and temporal dependencies, but computationally expensive and data-hungry.
- *IoT-based ML Models*: Improve real-time decision-making but require infrastructure investment.

Table 1. Comparison of machine learning and deep learning models for crop yield forecast.

Model	**Study**	**Best Performing Algorithm**	**R2 Score**	**Other Metrices (RMSE, M AE)**
Random Forest	Jhajharia et al. [2]	RF	0.963	0.035 (RMSE), 0.0251(MAE)
XGBoost	Maheswari & Ramasamy [4]	XGBoost	0.95	0.045 (RMSE)
CNN & Bi-GRU	Kolipaka & Namburu [3]	CNN & Bi-GRU	0.96	-
Extra Trees Regressor	Nikhil et al. [7]	Extra Trees Regressor	0.9615	21.06 (MAE), 33.99 (RMSE)
IoT-ML Integration	El basi et al. [5]	RF+IoT Sensors	0.95	-
Smart Farming AI	Patil et al. [6]	Naïve Bayes (for crop selection)	99.39% accuracy	-

CHALLENGES AND FUTURE DIRECTIONS

Key Challenges in ML-Based Crop Prediction

Data Accessibility and Quality

ML and DL models require large, diverse, and high-quality data sets. However, many agricultural datasets are incomplete, region-specific, or not publicly accessible [4, 6].

Environmental Variability

Crop yield depends on unpredictable climate conditions, pest infestations, and soil degradation, making it challenging to develop generalized ML models [3].

Computational and Infrastructure Costs

DL models require high computational resources, making them less feasible for small-scale farmers [7].

Incorporation with IoT and Remote Sensing

While IoT and remote sensing improve accuracy, implementation costs and technology adoption remain barriers in developing countries [5].

Future Research Directions

Hybrid models for better accuracy combining ML with remote sensing, IoT, and AI-driven automation can lead to more adaptive and accurate crop forecasting models [2, 6]. Development of lightweight AI models creating low-cost, resource-efficient DL models can enable real-time predictions for small- scale farmers [7–10]. Explainable AI (XAI) in agriculture future models should focus on transparency and interpretability, ensuring farmers and policymakers can understand and trust AI recommendations [3]. Climate-resilient AI models incorporating climate adaptation strategies into ML models can enhance crop yield prediction under extreme weather conditions [4].

CONCLUSION

The present review highlighted the advancements in ML and DL models for crop yield prediction, analyzing their performance, challenges, and future directions. While RF, XGBoost, CNN, and Bi-GRU remain among the most effective models, future research must address data accessibility, computational efficiency, and integration with IoT technologies. By focusing on hybrid AI approaches, climate resilience, and lightweight models, the agricultural sector can leverage AI-driven solutions for improved sustainability and food safety.

Acknowledgements

We are grateful to Dr. Vikas Gupta for providing valuable input and feedback through out there search procedure.

Funding

The authors received no monetary support for the research, authorship, and / or publication of the present article.

Conflict of interest

The authors declared no conflict of interest regarding this research. However, it should be noted that the first author of this paper is an employee of a Guru Kashi University Talwandi Sabo who is working as Dean of Department of Computing, Guru Kashi University Tawlandi Sabo. Second author is a research scholar, Department of Computing, Guru Kashi University Talwandi Sabo.

REFERENCES

1. Tiwari A, Darbari M. Emerging trends in computer science and its application. Proceedings of the International Conference on Advances in Emerging Trends in Computer Applications (ICAETC-2023); 2023 Dec 21–22; Lucknow, India.

2. Jhajharia K, Mathur P, Jain S, Nijhawan S. Crop Yield Prediction using Machine Learning and Deep Learning Techniques. Procedia Computer Science. 2023; 218: 406–417p.
3. Kolipaka VRR, Namburu A. Crop yield prediction using machine learning and deep learning techniques. International Journal of Recent and Innovation Trends in Computing and Communication. 2023; 11(10): 582–584p.
4. Maheswari MU, Ramasamy R. A comparative study of agricultural crop yield prediction using machine learning techniques. Proceedings of the 4th IEEE International Conference on Advanced Computing and Communication System; 2023 Mar 16–18; Shillong, Meghalaya.
5. Elbasietal E. Crop prediction model using machine learning algorithms. Applied Sciences. 2023; 13: 9288p.
6. Patil P, Athavale P, Bothara M, Tambolkar S, More A. Crop selection and yield prediction using machine learning approach. Current Agriculture Research Journal. 2023; 11(3): 968–980p.
7. Nikhil UV, Pandiyan AM, Raja SP, Stamenkovic Z. Machine learning-based crop yield prediction in South India: Performance analysis of various models. Computers. 2024; 13: 137p.
8. Petrillo A, Rehman M, Baffo I. Digital and sustainable transition in textile industry through internet of things technologies: A Pakistani case study. Applied Sciences. 2024; 14(13): 5380p.
9. Bukhari SNH. Data-Driven Farming-Harnessing the power of AI and machine learning in agriculture. USA: CRC Press; 2024.
10. Usang RO, Olu-Owolabi BI, Adebowale KO. Integrating principal component analysis, fuzzy inference systems, and advanced neural networks for enhanced estuarine water quality assessment. Journal of Hydrology: Regional Studies. 2025; 57: 102182p.

