

# Evaluating Effectiveness of Crop Yield Prediction using Machine Learning: A Comparative Analysis

Gitanjali Patil<sup>1</sup> and Hemant Kumar Gupta<sup>2</sup>

<sup>1-2</sup>SAGE University, Indore, India

Email: gitanjalipatil.in@gmail.com, hemugupta3131@gmail.com

**Abstract**— Precise forecasting of crop yields plays a pivotal role in advancing sustainable agricultural practices and ensuring global food security. This review paper presents a systematic evaluation of contemporary studies utilizing machine learning (ML) and deep learning (DL) techniques for yield prediction. The analysis encompasses investigations employing multi-modal data inputs, ranging from remote sensing data and soil characteristics to meteorological and climatic variables. The study critically examines the performance of various predictive models, including traditional machine learning approaches such as Linear Regression (LR), Support Vector Machines (SVM), Random Forest (RF), and Extra Trees (ET) as well as advanced neural network architectures such as Convolutional Neural Networks (CNN), Artificial Neural Networks (ANN), Deep Neural Networks (DNN), and Long Short-Term Memory networks (LSTM). Furthermore, the paper highlights current limitations in the field and suggests potential avenues for future research to enhance prediction accuracy and applicability.

**Keywords**— Machine learning (ML), Deep learning (DL), Linear Regression (LR), Support Vector Machines (SVM), Random Forest (RF), Convolutional Neural Networks (CNN), Long Short-Term Memory networks (LSTM).

## I. INTRODUCTION

India's economic framework has been fundamentally shaped by its agricultural sector, which continues to serve as a critical driver of both subsistence and national prosperity. For generations, this vital industry has provided economic stability for millions while simultaneously maintaining the country's nutritional requirements. Beyond its direct contributions, agriculture sustains an extensive network of secondary industries and forms the economic foundation of rural communities nationwide.

The sustainability of India's agricultural systems is increasingly jeopardized by the escalating impacts of climate change, which are fundamentally transforming conventional cultivation practices nationwide. Shifting climatic regimes characterized by inconsistent precipitation cycles and a rising frequency of severe weather phenomena are destabilizing agricultural productivity through altered growing seasons and compromised crop yields. This environmental volatility disproportionately affects marginal cultivators, whose limited access to capital and technology constrains their adaptive capacity. Furthermore, these climate-induced pressures intersect with chronic agricultural constraints, including deteriorating soil health, diminishing water resources, and evolving biotic stresses, collectively threatening the long-term viability of farming ecosystem.

India's farming systems face growing risks from unpredictable weather, which is disrupting crop production. Climate patterns have become more unstable, making traditional farming methods less effective. The monsoon, which farmers have relied on for generations, now behaves differently starting at uncertain times, spreading unevenly across regions, and ending unpredictably. These changes make it harder to grow crops successfully. This changing weather affects crop pollination, increases pests and diseases, damages soil health and it reducing harvests. These climate impacts worsen farmer debts and hunger in rural areas, creating serious economic and social challenges.

Modern agricultural research increasingly prioritizes yield forecasting through intelligent farming systems. Researchers integrate meteorological datasets, remote sensing outputs, and IoT-based field monitoring to generate precise crop production estimates. Such predictive analytics enable agricultural stakeholders to optimize cultivation strategies while assisting authorities in ensuring food security through data-driven resource allocation and risk mitigation protocols.

Accurate crop yield prediction faces significant challenges due to increasing climate variability. Erratic weather conditions, altered phenological cycles, and the rising frequency of extreme climatic events (e.g., unseasonal droughts, intense rainfall) disrupt conventional forecasting approaches. Current predictive models exhibit limited adaptability to these dynamic environmental shifts, necessitating the development of more resilient systems. Next-generation solutions must incorporate real-time data assimilation, adaptive learning mechanisms, and enhanced computational flexibility to deliver reliable agricultural projections under uncertain climatic regimes.

Modern yield prediction systems leverage machine learning (ML) and deep learning (DL) to extract actionable insights from multidimensional agricultural data, encompassing meteorological variables, edaphic properties, and multispectral remote sensing inputs. Conventional ML algorithms such as ensemble methods, support vector machines, random forest effectively identify critical yield determinants, while DL architectures such as convolutional and recurrent neural networks excel at modelling complex spatiotemporal relationships in high-dimensional datasets. Cloud-based infrastructures address computational constraints by offering distributed storage, parallel processing capabilities, and elastic resource allocation, enabling scalable agro-informatics solutions. This synergistic integration of computational paradigms enhances predictive performance under climatic uncertainty while supporting data-driven agricultural management through accessible, collaborative platforms.

Using Artificial intelligence (AI) for crop yield predictions faces several practical problems. First, these systems need huge amounts of detailed farm data that often doesn't exist for all areas or crop types. Second, the complex models of machine learning and deep learning that farmers can't understand how decisions are made, so they don't trust the results. Third, as climate patterns keep changing, the AI models need constant updates or they become less accurate over time.

The main objective of this paper is to perform comprehensive review of existed ML algorithm, DL algorithm for crop yield prediction. In this paper we study existed system for crop yield prediction and provide structural foundation for the future research.

## II. LITERATURE SURVEY

Sahoo, Singha, and Govind (2024) evaluated multiple machine learning algorithms including Cubist, random forest (RF), gradient boosting machine (GBM), multivariate adaptive regression splines (MARS), support vector machines (SVM), and extreme gradient boosting (XGB) to assess rice yield gaps in Eastern India under climate variability [1]. Their study incorporated climatic and soil variables, cross-validating model outputs against 1,889 field observations. Among the tested models, Cubist and RF demonstrated superior predictive accuracy, whereas the remaining algorithms exhibited limited performance. In a systematic literature review, Jabad and Murad (2024) examined the application of ML and DL techniques in crop yield forecasting. Their study evaluated the performance of algorithms such as RF, artificial neural networks (ANN), SVM, CNN, long short-term memory (LSTM) networks, and deep neural networks (DNN). The authors identified critical environmental predictors including precipitation, soil characteristics, and thermal conditions alongside vegetation indices (e.g., NDVI, EVI, LAI) as pivotal inputs for model training [2]. Noma and Babu (2024) proposed a machine learning framework based on Optimized Gradient Boosting to forecast the adoption of Climate-Smart Agriculture (CSA) practices. Utilizing longitudinal household survey data from Uganda's Rakai district, the researchers incorporated socio-economic indicators, agroecological conditions, and climatic variables as predictive features. Their model achieved a moderate prediction accuracy of 60%, demonstrating its potential in identifying CSA adoption trends [3]. Li et al. (2025) introduced a novel Knowledge-Guided Machine Learning (KGML) framework that incorporates outputs from the Agricultural Production Systems sIMulator (APSIM) to enhance

crop yield predictions under combined waterlogging and climate change conditions. Their methodology utilized extensive crop trial data from China's Yangtze River Delta region for model training and validation. The results demonstrated that the KGML approach significantly outperformed standalone APSIM simulations, achieving up to a 15% improvement in prediction accuracy ( $R^2$ ) [4].

Shawon et al. (2025) examined machine learning applications in crop yield prediction across studies published between 2017 and 2024. Following PRISMA methodology, the researchers analyzed 184 publications from multiple databases, ultimately selecting 97 studies for in-depth evaluation. Their analysis revealed that conventional machine learning approaches, particularly linear regression, random forests, and gradient boosted trees, remained prevalent, while convolutional neural networks and long short-term memory networks emerged as prominent deep learning techniques [5]. Sharma et al. (2024) evaluated the efficacy of three machine learning algorithms RF, Extra Trees (ET), and Artificial Neural Networks (ANN) for crop yield prediction. The study utilized an extensive dataset comprising 28,242 samples collected across 101 countries. Model performance was assessed using mean absolute error (MAE), coefficient of determination ( $R^2$ ), and mean squared error (MSE) metrics. Results demonstrated that the ET algorithm outperformed other models, achieving superior predictive accuracy of 97.5% with the lowest MAE value [6]. The SMART-CYPS framework developed by Kuradusenge and colleagues (2024) combines IoT sensors and machine learning to predict crop yields in sub-Saharan Africa. By monitoring weather and soil conditions in Rwanda's Musanze District, the system achieved MAPE values between 0.177-0.339 for potato and maize predictions. While demonstrating accurate forecasting capabilities, the authors recommend expanding the system's crop coverage and regional applicability to further support food security efforts [7]. Khanagoudar et al. (2023) performed a comprehensive systematic review examining machine learning applications for crop yield forecasting using climatic parameters. Their analysis evaluated multiple regression algorithms including decision trees, k-nearest neighbors, random forests, and linear regression models. The study incorporated critical predictive features such as meteorological conditions and plant pathology indicators [8]. study by Bloh, Lobell, and Asseng presented a knowledge-integrated machine learning system for wheat yield prediction, merging DSSAT simulations with neural networks and random forests. By generating synthetic training data from crop and climate models, their hybrid approach reduced prediction error by 8% versus standard methods, demonstrating the value of combining process-based modeling with machine learning for climate-resilient agriculture [9]. Jhajharia et al. (2023) investigated the application of machine learning and deep learning methodologies for agricultural yield forecasting across five major crops in Rajasthan, India. Their comparative analysis evaluated multiple predictive algorithms including Random Forest, Support Vector Machines, Gradient Descent optimization, Long Short-Term Memory networks, and Lasso regression. The Random Forest algorithm demonstrated superior performance, achieving a coefficient of determination ( $R^2$ ) of 0.963, along with low error metrics (RMSE = 0.035; MAE = 0.0251). The researchers highlighted the importance of incorporating diverse environmental variables to improve prediction accuracy and address critical challenges in agricultural yield estimation [10]. Pandya and Gontia (2023) conducted a comparative assessment of machine learning models for predicting cotton and groundnut yields in Gujarat's Saurashtra region, incorporating drought indices and remote sensing data. Their study evaluated multiple linear regression (MLR), multilayer perceptron-based artificial neural networks (ANN-MLP), and random forest (RF) algorithms. The RF approach demonstrated superior predictive accuracy, yielding  $R^2$  values ranging from 0.77 to 0.92, with root-mean-square errors (RMSE) of 80–133 kg/ha for cotton and 299–453 kg/ha for groundnut. The findings highlight the framework's potential for early drought detection and supporting agricultural planning [11].

Talaat, F. (2023) implements a novel Crop Yield Prediction Algorithm (CYPA) framework that integrates Internet of Things (IoT) technology with machine learning methodologies. The study comparatively evaluates five regression models Decision Tree, Random Forest, and Extra Trees regressors optimized through hyper parameter tuning. Experimental results demonstrated exceptional predictive performance, with model accuracy scores reaching 0.9814, 0.9933, and 0.9903 respectively. Furthermore, the authors incorporated an active learning component to the CYPA system, which significantly enhanced both computational efficiency and prediction accuracy in agricultural yield forecasting applications [12]. Srivastava et al. (2022) proposed a 1D-CNN model for predicting winter wheat yields in 271 German counties over a period from 1999 to 2019. The model incorporated data on weather, soil conditions, and phenology. It outperformed eight baseline models, including RF, XGBoost, and DNN, by reducing RMSE by 7–14% and enhancing correlation by 4–50%. SHAP analysis revealed that factors such as DUL, wind speed during week 10, and radiation in week 7 played a significant role in predictions. Although the model demonstrated high accuracy, its black-box nature and the regional variability of errors were noted as limitations [13]. Van Klompenburg et al. (2020) conducted a Systematic Literature Review (SLR) to examine crop yield prediction using machine learning. They retrieved

567 studies from six databases, selecting 50 for in-depth analysis based on inclusion/exclusion criteria. Key features (temperature, rainfall, soil type) and algorithms (e.g., Artificial Neural Networks) were extracted. A supplementary search identified 30 deep learning studies, analyzing CNN, LSTM, and DNN [14]. Dilli Paudel and others combines agronomic principles with machine learning techniques to develop a flexible and reusable framework for crop yield prediction. By leveraging inputs such as crop simulation data, meteorological conditions, soil properties, and satellite-derived remote sensing metrics, the model generates yield forecasts for five distinct crops across three different countries, utilizing the MCYFS database. When evaluated against MCYFS benchmark results, the proposed system demonstrates comparable or enhanced accuracy for specific crops and geographic areas. The inclusion of remote sensing data further refines predictive performance. Findings suggest that machine learning-based approaches can effectively emulate—and in certain cases, surpass—conventional yield forecasting methods, providing advantages in adaptability, precision, and scalability for regional or global agricultural assessments[15]. Michele Meroni and others presents a machine learning (ML) framework designed for crop yield prediction under data constraints, with a case study in Algeria. The study identifies key predictive features and evaluates model performance for barley, soft wheat, and durum wheat yield forecasting. Despite limited dataset sizes, the ML approach achieved strong predictive accuracy, with mean yield errors ranging between 13% and 14%. Comparative analysis revealed that ML models consistently surpassed simpler benchmark models, particularly in low-yield years, though practical differences were occasionally marginal. Approximately 60% of the ML implementations outperformed baseline models, underscoring the critical role of model calibration and selection. The results demonstrate that machine learning remains a viable tool for agricultural yield prediction, even in data-scarce scenarios [16].

### III. TECHNOLOGIES USED FOR CROP YIELD PEDICTION

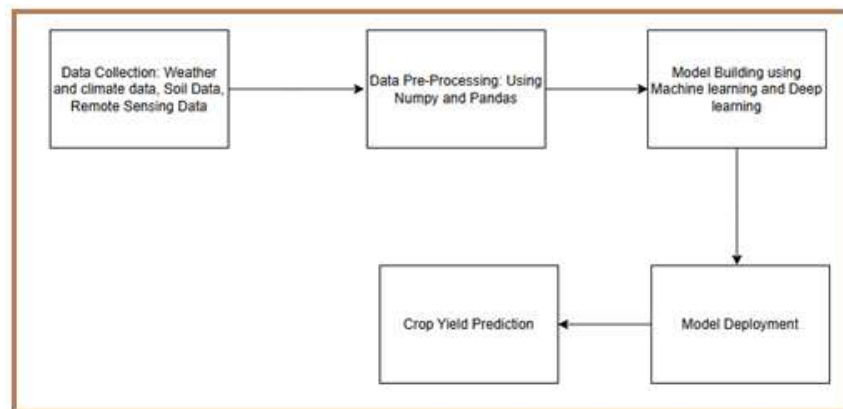


Figure 1: Crop yield prediction using ML process

This above figure 1 demonstrate a comprehensive framework for crop yield prediction by leveraging heterogeneous data sources, including meteorological records, soil properties, and remote sensing imagery. The pre-processing phase utilizes Python libraries such as NumPy and Pandas to clean and normalize datasets, ensuring reliable input for predictive modeling. For accurate yield estimation, the system employs both traditional machine learning algorithms (e.g., Random Forest for feature importance analysis) and advanced deep learning architectures, including convolutional neural networks (CNNs) for spatial data processing and long short-term memory (LSTM) networks for temporal pattern recognition. Although machine learning techniques provide greater interpretability, deep learning proves more effective in analysing high-resolution satellite imagery and large-scale environmental datasets. Key limitations include inconsistencies in data formats, computational demands for large-area deployment, and the necessity for field-level validation to ensure practical relevance.

To facilitate adoption, the proposed model is designed for deployment through accessible digital platforms, such as mobile applications or cloud-based dashboards, enabling farmers to make data-driven decisions. Future enhancements could incorporate IoT-based sensor networks and edge computing infrastructure to support dynamic, real-time yield forecasting. While this approach highlights the transformative potential of AI in agriculture, further refinement is needed to mitigate dataset biases and improve model resilience against climate anomalies, ensuring scalability across diverse agro ecological regions.

#### IV. COMPARATIVE ANALYSIS

This section presents a comparative analysis of the literature reviewed by various authors. A detailed summary of the methodologies and training approaches used by the researchers is provided in the following table.

TABLE I: COMPARATIVE ANALYSIS

| Authors and their work                                                                                                                                                                                          | Methodology and Dataset Used                                                                                                                                                                                                                                | Analysis                                                                                                                                                                                                                      | Limitations                                                                                                                             |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| Sahoo, Singha, and Govind (2024) focused on predicting rice yield gaps in Eastern India under climate uncertainty using machine learning models like Cubist, RF, GBM, MARS, SVM, and XGB.                       | Climate scenarios were generated using CMIP6 data (SSP2-4.5 and SSP5-8.5). Model validation was performed using 1889 field-based farmer observation records.                                                                                                | Performance was evaluated through $R^2$ values. Cubist ( $R^2 = 0.73$ ) and RF ( $R^2 = 0.72$ ) models performed best. Other models like XGB, GBM, SVM, and MARS showed poor predictive accuracy                              | Some machine learning models such as XGB, GBM, SVM, MARS showed low predictive performance.                                             |
| Md. Abu Jabeed and Masrah Azrifah Azmi Murad (2024) conducted a comprehensive review of AI-based methods (ML and DL) for crop yield prediction [2].                                                             | No original dataset used. They used MODIS satellite data, Landsat images, soil databases, and weather datasets                                                                                                                                              | Explored various ML and DL models such as RF, ANN, SVM, CNN, LSTM, and DNN for crop yield estimation. Emphasized the critical role of accurate environmental data and model selection for improving yield prediction accuracy | They perform only review based study. Dynamic nature of the environmental factors limit model generalizability across different regions |
| Freddy Noma and Suresh Babu (2024) explored predicting CSA practices using an optimized Gradient Boosting ML model trained on farmers' household-level data from Uganda [3].                                    | Used longitudinal survey data from Rakai district, Uganda. Variables included socio-economic characteristics, farm features, agroecological features, and climate indicators (temperature, rainfall). Optimized Gradient Boosting ML algorithm was applied. | Two CSA response variables were predicted separately. The model achieved fair prediction accuracy, with scores up to 60%.                                                                                                     | Limited predictive accuracy nearly to 60%. Need to use more comprehensive dataset for better accuracy.                                  |
| Li, He, Harrison, Shi, Feng, Wang and others (2025) developed a Knowledge-Guided ML (KGML) model to improve crop yield projections under waterlogging conditions affected by climate change [4].                | Integrated APSIM crop model outputs with machine learning models trained on experimental crop trial data from the Yangtze River Delta, China. Climate scenarios were used to simulate future waterlogging impacts.                                          | KGML models consistently improved predictive accuracy compared to APSIM-only simulations, with up to 15% higher $R^2$ values.                                                                                                 | Need a quality dataset for getting a better result. Further research is needed for wider region and other crop types.                   |
| Sarowar Morshed Shawon and others performed systematic literature review on ML and DL for crop yield prediction (2017–2024) [5].                                                                                | Systematic Literature Review (SLR) using PRISMA technique; 184 papers from 8 electronic sources filtered to 97 papers.                                                                                                                                      | They used features such as Temperature, soil type, vegetation. Used ML model such as LR, RF, GBT and used DL model such as CNN, LSTM.                                                                                         | Limited Data diversity for model training. It need for real-world validation of hybrid models.                                          |
| Sharma and colleagues (2024) performed a machine learning comparison study for agricultural yield forecasting [6].                                                                                              | Performed Comparative evaluation using 28,242 samples from 101 countries. Build ML and DL model such as RF, ET and ANN.                                                                                                                                     | ET model performed best with 97.5% accuracy and ANN showed lower reliability                                                                                                                                                  | They used traditional machine learning model. Need advanced algorithms and larger datasets for better validation.                       |
| Martin Kuradusenge and others introduced SMART-CYPS, an innovative framework integrating Internet of Things (IoT) technology and machine learning for agricultural yield forecasting in sub-Saharan Africa [7]. | Used IoT sensor data of rainfall, temperature, humidity, soil moisture. They focus on Irish Potato and Maize over multiple agricultural seasons.                                                                                                            | SMART-CYPS system showed high predictive accuracy for food security improvement.                                                                                                                                              | Limited to specific crops such as Irish potato, maize and one region. Scalability to other crops and regions not yet tested.            |
| Khanagoudar and colleagues analysed climate-based yield prediction using machine learning [8].                                                                                                                  | They used features such as weather, plant diseases and crop classification. Implement ML techniques such as Decision Tree Regressor, KNN Regressor, Random Forest Regressor and Linear Regressor                                                            | They highlights neural networks struggles with non-linear input-output relationships. Discusses strengths and weaknesses of different ML models                                                                               | Neural networks face difficulty in non-linear mapping. Need to improve quality of the dataset.                                          |
| Malte von Bloh, David Lobell, and others (2024) proposed a hybrid machine learning                                                                                                                              | They implemented a hybrid modelling framework that integrates the process-based DSSAT                                                                                                                                                                       | Hybrid model improved prediction accuracy by 8% over purely data-driven models.                                                                                                                                               | Depends heavily on quality and assumptions of                                                                                           |

|                                                                                                                                                          |                                                                                                                                                                                                        |                                                                                                                                                                                        |                                                                                                                                          |
|----------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------|
| approach combining domain knowledge from process-based models (DSSAT) with neural networks [9].                                                          | (Decision Support System for Agro technology Transfer) simulations with machine learning techniques, specifically neural networks and random forest algorithms.                                        | synthetic samples boosted performance of the model                                                                                                                                     | synthetic data. Need to adapt broader crop scenarios.                                                                                    |
| Kavita Jhajharia and others (2023) explored machine learning and deep learning techniques to predict crop yield in Rajasthan using five crop types [10]. | Applied multiple ML and DL models such as Random Forest, SVM, Gradient Descent, LSTM, Lasso Regression.                                                                                                | Random Forest performed best with 0.963 R <sup>2</sup> , 0.035 RMSE, and 0.0251 MAE. They Highlighted importance of integrating multiple environmental factors.                        | They limited only for specific geographical location. Further expansion to different regions is needed.                                  |
| Parthasarathi Pandya and others (2023) evaluated crop yield prediction models for cotton and groundnut in Gujarat's Saurashtra region [11].              | Used multiple linear regression (MLR), artificial neural networks (MLP-ANN), and random forest (RF) with meteorological and Landsat vegetation indices from 1986–2019 for Gujarat's Saurashtra region. | RF models performed best accuracy with R <sup>2</sup> values ranged from 0.77–0.92, high prediction efficiency for cotton and groundnut yields                                         | Their research is only limited to Saurashtra region May need model adjustment and validation for different crops and geographic regions. |
| Talaat, F. M. (2023) employs the Crop Yield Prediction Algorithm (CYPA) utilizing IoT techniques to predict crop yields [12].                            | Utilized IoT techniques and trained five machine learning models with optimal hyper-parameter settings.                                                                                                | Decision Tree Regressor achieved accuracy of 98%. Random forest regressor achieved accuracy of 99% and Extra tree regressor achieved accuracy of 99%.                                  | It will reliance on specific environmental variables. They used limited labelled data for training.                                      |
| Srivastava et al. (2022) Proposed a 1D-CNN for winter wheat yield prediction in Germany [13].                                                            | Use weather data having feature such as temperature, radiation, precipitation. Also use soil and phenology data from 271 counties. They implemented model such as 1D-CNN, RF, XGBoost and DNN.         | CNN achieved lowest RMSE and highest correlation ( $r = 0.65$ to $0.81$ ). SHAP revealed DUL and wind speed as top features.                                                           | Dependency on high-quality and long-term data for better accuracy.                                                                       |
| van Klompenburg et al Conducted an SLR on crop yield prediction using machine learning and deep learning techniques [14].                                | They Analysed 50 ML and 30 DL studies; common datasets included temperature, rainfall, soil, and satellite imagery                                                                                     | They used features such as Temperature, rainfall, and soil type. They analyse model such as CNN, LSTM and DNN.                                                                         | Findings may vary based on regional data and crop types                                                                                  |
| Dilli Paudel and others Combined agronomic modeling principles with machine learning for large-scale crop yield forecasting [15].                        | Used a modular and explainable ML framework. Data from MCYFS uses crop model outputs, weather, remote sensing, soil data.                                                                              | Compared predictions to MCYFS operational forecasts. ML models matched or exceeded MCYFS accuracy for several crops and conditions, particularly with added remote sensing indicators. | Performance varied by crop and country. In some cases, MCYFS still outperformed the ML models. Further optimization needed               |
| Michele Meroni and others Developed ML models to forecast yields using small data in Algeria [16].                                                       | Built a machine learning pipeline selecting best features and models. Used climate and remote sensing data for barley, soft wheat, and durum wheat yield predictions.                                  | ML models outperformed benchmark models in 60% of cases. Achieved 13–14% mean error rates within season. Confirmed strong performance especially in low-yield seasons.                 | Differences between ML and benchmarks were sometimes minor. ML superiority required careful model calibration and data selection.        |

## V. CONCLUSION

This comprehensive review demonstrates that advanced computational methods, particularly machine learning and deep learning approaches, hold considerable promise for enhancing the precision and effectiveness of crop yield forecasting systems.

The synthesis of multiple data modalities - from spectral satellite observations to edaphic factors and meteorological variables - has enabled the creation of increasingly refined predictive models.

Conventional machine learning algorithms continue to serve as reliable baseline methods, while contemporary deep learning frameworks show particular efficacy in processing voluminous, multi-dimensional agricultural datasets.

Nevertheless, persistent limitations regarding data integrity, model transferability across regions, and computational scalability must be addressed. Subsequent investigations should prioritize the development of more resilient and transparent modeling frameworks, with particular emphasis on integrated methodologies that combine the strengths of different algorithmic approaches to optimize predictive performance in real-world agricultural applications.

## REFERENCES

- [1] Sahoo, S., Singha, C., & Govind, A. (2024). Advanced prediction of rice yield gaps under climate uncertainty using machine learning techniques in Eastern India. *Journal of Agriculture and Food Research*, 18, 101424
- [2] Jabeed, M. A., & Murad, M. A. A. (2024). Crop yield prediction in agriculture: A comprehensive review of machine learning and deep learning approaches, with insights for future research and sustainability. *Heliyon*, 10, e04836.
- [3] Noma, F., & Babu, S. (2024). Predicting climate smart agriculture (CSA) practices using machine learning: A prime exploratory survey. *Climate Services*, 34, 100484
- [4] Li, L., He, Q., Harrison, M. T., Shi, Y., Feng, P., Wang, B., Zhang, Y., Li, Y., Liu, D. L., Yang, G., Zhou, M., Yu, Q., & Liu, K. (2025). Knowledge-guided machine learning for improving crop yield projections of waterlogging effects under climate change. *Resources, Environment and Sustainability*, 19, 100185.
- [5] Shawon, S. M., Ema, F. B., Mahi, A. K., Niha, F. L., & Zubair, H. T. (2025). Crop yield prediction using machine learning: An extensive and systematic literature review. *Smart Agricultural Technology*, 10, 100718.
- [6] Sharma, S., Walia, G. K., Singh, K., Batra, V., Sekhon, A. K., Kumar, A., Rawal, K., & Ghai, D. (2024). Comparative analysis on crop yield forecasting using machine learning techniques. *Rural Sustainability Research*, 52(347), 2024
- [7] Kuradusenge, M., Hitimana, E., Mtonga, K., Gatera, A., Habiyaemye, J., Ngabonziza, J., Hanyurwimfura, D., Rukundo, P., & Mukasine, A. (2024). SMART-CYPS: An intelligent internet of things and machine learning powered crop yield prediction system for food security. *Discover Internet of Things*.
- [8] Khanagoudar, P. S., Sushma, B. S., Chandana, C. N., Manikanta, N., & Awatimath, S. S. (2023). Crop yield prediction using machine learning algorithm based on climate variables. *International Journal of Cheminformatics*, 1(2).
- [9] von Bloh M., Lobell, D., & Asseng, S. (2024). Knowledge informed hybrid machine learning in agricultural yield prediction. *Computers and Electronics in Agriculture*, 227, 109606.
- [10] Jhajharia K., Mathur, P., Jain, S., & Nijhawan, S. (2023). Crop yield prediction using machine learning and deep learning techniques. *Procedia Computer Science*, 218, 406–417.
- [11] Pandya P., & Gontia N. K. (2023). Early crop yield prediction for agricultural drought monitoring using drought indices, remote sensing, and machine learning techniques. *Journal of Water and Climate Change*, 14(12), 4729–4743.
- [12] Talaat F. M. (2023). Crop yield prediction algorithm (CYPA) in precision agriculture based on IoT techniques and climate changes. *Neural Computing and Applications*, 35, 17281–17292
- [13] Srivastava A. K., Safaei N., Khaki S., Lopez G., Zeng W., Ewert F., Gaiser T., & Rahimi, J. (2022). Winter wheat yield prediction using convolutional neural networks from environmental and phenological data. *Scientific Reports*, 12(1), 32315.
- [14] van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. *Computers and Electronics in Agriculture*, 177, 105709.
- [15] Paudel D., Boogaard, H., de Wit A., Janssen S., Osinga, S., Pyliannidis C., & Athanasiadis I. N. (2021). Machine learning for large-scale crop yield forecasting. *Agricultural Systems*, 187, 103016.
- [16] Meroni, M., Waldner, F., Seguíni, L., Kerdiles, H., & Rembold, F. (2021). Yield forecasting with machine learning and small data: What gains for grains? *Agricultural and Forest Meteorology*, 308-309, 108555