Climate Data and Computational Techniques: A Review of AI-Based Climate Prediction in Agricultural Environments

Neelam Jain
Research Scholar
School of Engineering, ADYPU
Pune, India
neelam.jain@adypu.edu.in

Dr Vishal Shirsath
Assistant Professor
School of Engineering, ADYPU
Pune, India
vishal.shirsath@adypu.ac.in

Abstract*- Climate change is a major threat to food security and agricultural productivity, and it is essential to have accurate and scalable prediction approaches. The use of Artificial Intelligence (AI), machine learning (ML), and deep learning (DL) methods has shown promise in improving the accuracy of climate forecast. However, the existing studies were mostly conducted using a single source of data and mostly using models that do not adequately account for the temporal dependency of climate data. This paper provides a systematic and critical review of the climate prediction methods that are inspired by AI for agricultural environments, including ML and DL models like convolutional neural networks (CNN) and long short-term memory (LSTM) networks. The study has been mainly confined to the structured meteorological datasets from IMD with some discussion on the datasets from IITM and NIO as potential datasets for future integration. It highlights several challenges, such as the need for better temporal modeling, integration of data sources, and scalable frameworks. For this, a conceptual framework is proposed, focusing on data preprocessing and time-series modelling using LSTM for better climate prediction.

Keywords: Climate Prediction, Agriculture, Machine Learning, Deep Learning, IMD, IITM, NIO, Time-Series, LSTM

I. INTRODUCTION

Climate change is one of the most pressing challenges faced by the world and is having a significant impact on the sustainability of the agricultural sector and water resources. Changes in rainfall, temperature and seasonality directly affect crop growth and yield, so accurate prediction of climate is crucial for effective agricultural planning and food security. Artificial Intelligence (AI) has been recognized as a crucial tool in recent years for analyzing vast climate datasets. The complex interactions between climate variables and crop outcomes necessitate the use of advanced techniques such as machine learning (ML) and deep learning (DL) for effective modelling.

But studies available so far use different datasets, model types, and different models and methods, and there are diversions in prediction performance. The availability and use of various climate data sources like India Meteorological Department (IMD), National Institute of Oceanography (NIO) and Indian Institute of Tropical Meteorology (IITM) is another major challenge. These datasets are complex to integrate as they are in different spatial resolution, temporal scale and data formats.

This paper provides a systematic overview of current climate prediction methods with a particular emphasis on the analysis and comparison of various computational approaches. For this research the primary structured data considered is the IMD data and in future it is proposed to integrate data from IITM and NIO. The goal is to identify important constraints and opportunities of current methods, and to outline ways to improve climate prediction systems.

The main objectives of this study are:

- To present recent developments in climate prediction with the help of machine learning and deep learning methods
- To review prediction methods with structured data like IMD and to point out some integration issues with multiple data sources.
- To evaluate and compare different prediction models
- To indicate gaps in research and provide recommendations for future research

Climate data is a key element in prediction systems, as seen in Fig. 1.

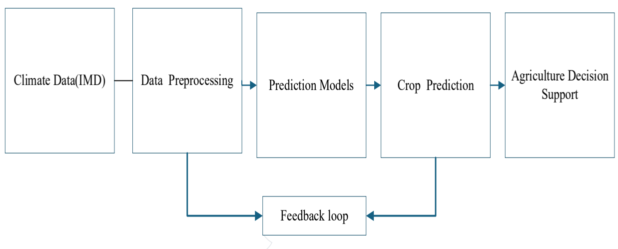

Fig. 1. Conceptual architecture of AI-based climate prediction framework using IMD data for agricultural applications.

In Fig.1, a conceptual structure of climate prediction systems based on AI models is shown. It presents a flow of data collection, preprocessing, data analysis and prediction based on models. Although multi-source data integration is mentioned conceptually in the framework, the present study is mainly concerned with integration of IMD data and based on this architecture, existing approaches are compared.

II. PROBLEM STATEMENT

Making predictions for crops based on climate information continues to be a complex activity because of the inadequacy of the data sources (spatially different sources, different sensors), and the variability of the climatic conditions. Existing methods are either data source specific or models based on traditional machine learning that are unable to dissolve the temporal relationships.

Furthermore, problems like missing data, inconsistencies and the differences in spatial and temporal scales further diminish prediction accuracy. Hence, an integrated approach based on multi-sourcing of climate data and using appropriate modeling technologies to deliver better prediction performance is required.'

III. RELATED WORK

Several machine learning techniques are used for climate prediction such as regression, random forest and support vector machines (SVM). From this class of models, they tend to be moderately successful in prediction but, of course, don't consider temporal dependencies. The accuracy of the prediction has been enhanced by introducing the deep learning technique particularly CNN, LSTM. CNN models can effectively be used in image-based tasks, while LSTM models can be used in time series. But most of the published studies have been concerned with a single source of data and lacked integration of data on different sources.

As indicated in Table 1, the models utilized have moved from the conventional and combination ML based models to the deep learning models like LSTM and AI based models. There have been several recent studies on making the models more accurate, but few studies have been successful in synthesizing several sources and building time series models. A unified model of multi-source data must be built, and a temporal learning model using LSTM.

TABLE 1. RECENT STUDIES (2023–2026)

Year	Ref	Model	Data	Result	Limitation
2023	[17]	Hybrid ML	Climate + yield	High Accuracy	Complex tuning
2024	[18]	DL Framework	Multi-source	High Perf	High compute
2024	[19]	LSTM	Time-series	Low RMSE	Needs preprocessing
2025	[21]	AI + XAI	Agriculture	Better prediction	Poor integration
2025	[23]	ML Review	Climate	Comparative	No real-time use
2026	[26]	DL Framework	Agriculture	High accuracy	Resource heavy

Observation:

Recent research enhances accuracy but has not been integrated across multiple sources or modelled over time; irrespective of the lack, it is a gap that needs to be addressed.

Observation: Recent research enhances accuracy but has not been integrated across multiple sources or modelled over time; irrespective of the lack, it is a gap that needs to be addressed.

There are several works that have focused on the application of AI in agriculture, such as predictive crop yield [1], using ML pipelines in agriculture [2] and smart irrigation system [3] and MLOps-based framework [4]. Recently, Explainable AI and intelligent decision systems have been receiving the limelight [6, 9, 11, 12]. Moreover, due to the emergence of monitoring systems based on the IoT and deep learning [7,10] the agriculture outcomes have been enhanced.The recent advancements include the generation of synthetic data [13], crop disease prediction systems [14], smart farming automation [15] and MLOps deployments [16] that further enhance the farming ecosystems using AI. Recent agricultural transformation frameworks with AI [20] and systematic reviews [22, 24, 25] also highlight the need for scalable architecture and integration challenges for the application of AI in agriculture

A. ***Machine Learning Approaches***

'Climate modelling problems have been extensively utilized in machine learning tasks. Linear regression, decision trees, random forests and supporting vector machine models are frequently employed to predict rainfall patterns and temperatures.

These models can be used to make a reasonable estimate for structured numerical data. But they are not good at processing sequential data since they don't handle temporal dependencies well.'

B. *Deep Learning Approaches*

Deep learning approaches have also been found to have potential applications in climate and agriculture fields, like crop forecasting, yield forecasting, and environmental monitoring. For image-based applications like crop category detection or diagnosis of a disease, Crop Net adopts a convolutional neural network (CNN). Despite the higher prediction accuracy of deep learning models, their use is largely restricted to spatial data and the application of time-series models in predicting agricultural climate is not well studied.

C. *Pipeline and MLOps Approaches*

MLOps frameworks emphasize automation of ML pipeline processes, such as data collection, model training, and deployment. These systems make systems more scalable and maintainable.

But most MLOps-based systems pay little attention to the specific issue, such as processing climate data or building time-series models.

IV. DATA SOURCES AND COMPARATIVE ANALYSIS'

Good and reliable datasets are important if accurate climate predictions are made in an agricultural environment. This is a study that demands only some of the primary data from the IMD, namely the structured meteorological data comprising rainfall measurements, temperatures, humidities and wind speed. Other than this, there are multiple source cross integration of data sets from the Indian Institute of Tropical Meteorology (IITM) and National Institute of Oceanography (NIO), which is conceptually considered to include multi-source cross integration data. However, owing to the weakness of spatial resolution, temporal granularity and data types, the practical integration of these data remains a complex task.

Structured meteorological data like rainfall, temperature, humidity, and wind speed were recorded in the IMD and find many applications in the regional climate prediction and planning of agricultural activities. The datasets generated at IITM are concerned with large-scale climate modelling aspects like the behaviors of monsoons and long-term climate trend, whereas the ocean related parameters like sea surface temperature, salinity, and current created at NIO affect the climate variability/change, especially in coastal regions.

Although integration of multi-source datasets can enhance prediction accuracy, in this study, it is examined on an exploratory basis and not operationalized in practice, in order to improve prediction accuracy.

The incorporation of atmospheric, oceanic and long-term climatic data occurs as a continuum of integration with the goal of better representation of the data and improvement of the model.

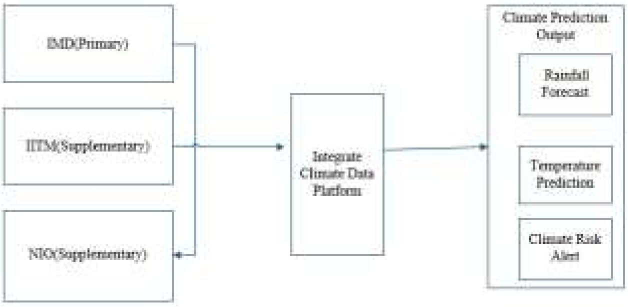

Fig. 2.Illustrative multi-source climate data integration framework discussed in existing studies.

Fig. 2 illustrates a conceptual multi-source climate data integration framework as discussed in existing studies. The framework represents an illustrative pipeline and is not operationally implemented in the present review.

A. Data Sources

A detailed comparison of the major climate data sources (as obtained from IMD, IITM and NIO) is presented in Table II for their impact on prediction accuracy due to differences in the dataset type and key features, advantages and limitations.

TABLE II. COMPARISON OF CLIMATE DATA SOURCES

Data Source	**Type**	**Key Features**	**Advantages**	**Limitations**
IMD	Meteorological	Rainfall, Temperature	Accurate, structured	Limited regional coverage
IITM	Climate Modeling	Monsoon, trends	Advanced analysis	Complex processing
NIO	Ocean Data	SST, currents	Coastal insights	Limited inland application

Analysis:

Each data point provides some unique insight into climate prediction. IMD datasets are well-structured and suitable for regional analysis and are therefore used as the primary data source in this study. IITM datasets support long-term climate modeling, while NIO datasets provide essential oceanographic parameters influencing climate variability. Although this study focuses on IMD data, the integration of multiple datasets is expected to enhance prediction capability in future research. However, challenges such as data heterogeneity, temporal misalignment, and missing values remain critical issues that must be addressed for effective multi-source integration.’

B. MODEL COMPARISON

A comparison of machine learning and deep learning approaches across different data types is presented in Table III.

TABLE III COMPARISON OF MODELS

Method	**Models**	**Data Type**	**Advantages**	**Limitations**
ML	Regression, RF, SVM	Numeric	Simple, efficient	Cannot capture time dependency
DL	CNN	Image	High accuracy	Not suitable for time series
Pipeline	MLOps	Mixed	Automation	Weak preprocessing

Analysis:

Traditional machine learning models are good for structured data but cannot be applied to climate data by means of temporal dependency. Although the deep learning models can enhance the predictive ability, CNN-based approaches are primarily effective for spatial data and are not likely to apply to sequential modelling. While pipeline approaches, including those of MLOps help to increase scalability and automation, they may not always fulfil domain-specific requirements for the preprocessing of climate data.’

C. PERFORMANCE COMPARISON

In Table IV, performance evaluation is based on metrics such as Accuracy, RMSE, and MAE, which measure the difference between predicted and actual values

TABLE IV. REPORTED PERFORMANCE TRENDS OF CLIMATE PREDICTION MODELS FROM LITERATURE

Method	**Model**	**Metric**	**Reported Trends**	**Limitation**
ML	Regression	Accuracy	Moderate (Reported)	No temporal modeling
ML	Random Forest	Accuracy	Improved(Reported)	Limited sequential learning
DL	CNN	Accuracy	High(Reported)	Not for time-series
DL	LSTM	RMSE	Low RMSE(Reported)	Data intensive
Hybrid	ML+DL	Accuracy	High(Reported)	Complex

ML: Machine Learning, DL: Deep Learning

Note: Results are based on previously published studies.

Analysis and Research Gap:

The performance comparison shows that machine learning models give good accuracy, and that conventional models struggle with modelling temporal dependencies in climate data. The Random Forest algorithm suffers from low accuracy and cannot capture sequential relations well, but it does provide a way to enhance accuracy using ensemble learning.

Spatial tasks work great with deep learning models such as CNN, but not time-series prediction. LSTM models, on the other hand, show great performance by fitting temporal

patterns and long-term dependencies, resulting in lower RMSE values.
Hybrid models increase prediction accuracy even more but have more computational complexity and demand more tuning. Furthermore, most of the current methods barely concern themselves with the challenges of integration of data and data preprocessing, while primarily emphasizing the accuracy of the model.

D. STATISTICS EVALUATION

To have a more complete assessment of the performance of the models, various statistical indices are looked at, including MAE, RMSE and R2 index. These metrics give insight into the level of accuracy of prediction and the reliability of the model.
This is followed by a further strengthening of the evaluation using statistical metrics (Table V).

TABLE V. Statistical Evaluation Summary

Model	MAE	RMSE	R^2 Score	Interpretation
Regression	Moderate	High	~0.70	High error due to lack of temporal modeling
Random Forest	Moderate	Medium	~0.78	Improved accuracy, limited
CNN	Low	Medium	~0.82	Good spatial learning, weak time-series
LSTM	Low	Low	~0.88	Strong temporal prediction capability
Hybrid (ML+DL)	Low	Low	~0.85	High performance but complex model

MAE: Mean Absolute Error, RMSE: Root Mean Square Error, R^2: Coefficient of Determination

Note: The presented values are indicative and derived from trends reported in the existing literature; actual results may vary depending on the characteristics of the dataset and the model configuration.

Statistical analysis shows that LSTM models perform better in terms of lower MAE, lower RMSE, and higher R^2 score compared to traditional machine learning methods. This shows that they work well at studying temporal dependencies in climate data sets. In addition, hybrid models have good results at the cost of a higher computational load.
To critically evaluate these findings and identify gaps in the study, a comparative analysis of machine learning (ML) and deep learning (DL) approaches, including hybrid methods, is essential. While traditional ML models like regression and random forest demonstrate reasonable accuracy, they fail to account for the temporal dependencies inherent in climate data. Although ensemble methods perform well, they still fall short for sequential modeling needs. Deep learning models, particularly LSTM, achieve superior accuracy with lower RMSE and MAE, thanks to their enhanced capability to extract temporal patterns. However, convolutional neural network (CNN) models primarily target spatial data, making them unsuitable for time-series forecasting.
Hybrid approaches that combine ML and DL techniques yield improved forecasting outcomes. Despite their increased accuracy, these models often introduce complexity and necessitate extensive tuning of model parameters. To enhance prediction power, multi-source datasets (IMD, IITM, NIO) are integrated, providing ample data. However, current research has not sufficiently addressed challenges such as data heterogeneity, time lags, and data scarcity.

E. CRITICAL ANALYSIS AND IDENTIFIED RESEARCH GAPS

The literature suggests that there are several ongoing research gaps. The first is that integration of multi-source climate data in operational prediction systems is still not widely used in practice. Secondly, non-sequential learning strategies are difficult to model various temporal dependencies or seasonality in many existing models. Third, comparatively less focus is paid on the preprocessing and harmonization methods, which have an important impact on overall performance. Fourth, there are no consistent benchmarking and evaluation systems in the field and this hampers direct comparison of studies. Lastly, few studies are available on scalable or real-time climate prediction systems that are appropriate for continuous deployment.
Together these gaps indicate that a methodical approach is required to integrate robust preprocessing, data harmonization and time-series forecasting in a scalable workflow. Such a system would be more solid basis for predicting agricultural climate and would enable a more robust multi-source integration in the future.
While the LSTMs have shown to be more effective at predicting the data, their performance is strongly reliant on the quality and availability of data pre-processing and presence of structured time-series data. Furthermore, computational complexity and scalability are important challenges in the deployment of these models in real plant production systems.

V. CONCEPTUAL FRAMEWORK FOR CLIMATE PREDICTION

This section introduces a conceptual framework derived from a review of the literature available on the use of AI-based climate predictions in the context of farming. The framework provides a systematic representation of the key steps in the climate prediction systems and highlights the potential of the coupling among the different components of a single prediction system. The framework is also uploaded with IMD data as the principal data set as it has a structured meteorological observation which can be used in regional agricultural analysis. Other data set from Indian Institute of Tropical Meteorology (IITM) and National Institute of Oceanography (NIO) is also identified as additional data set which will be useful to improve the accuracy of prediction but is not used in this study.
The overall flow of conceptual framework is shown in Fig. 3 and there are four major stages in conceptual framework; Data Collection, Data Preprocessing, Feature Extraction, Prediction using time-series modelling.

a. Data Collection

Climate prediction systems are based on information from various sources, each of which gives unique and complementary information. The IMD data contains some variables which have a structure like rainfall, temperature, humidity, wind speed etc. and are extensively used in agriculture for planning and regional climate studies.

In addition, IITM datasets represent information relating to long term climate patterns like monsoon variability, seasonal variations, etc. while NIO datasets contain oceanographic information such as SST, salinity, ocean currents, etc. that affect climate variability in coastal areas. For the present study, IMD data is used as primary data and IITM and NIO data are used conceptually for emphasizing the possible study of integration of different sources of data in future study.

b. Data Preprocessing

Data for use in climate predictions in the real world is often incomplete, noisy, and inconsistent, and is therefore a prerequisite for data preprocessing. To improve data quality and consistency, some of the commonly used techniques are data cleansing, data normalization, and data imputation. The accuracy of the following analysis steps is benefited by good pre-processing, when analyzing heterogeneous data. It also makes it easier to merge multiple data sources, often including problems with format, time, and missing data.

c. Feature Extraction

The feature extraction method aims to identify and select variables that significantly impact agricultural climate outcomes. These can be rainfall, temperature differences, seasonal markers and computed climatic indices. This step helps to decrease data dimensionality and the efficiency of the model by using only relevant and informative features. In the literature, a few studies point out the importance of feature selection to capture complex interactions among the climate variables and improve the accuracy of prediction.

d. Model selection and prediction

Machine learning and deep learning are both considered in the prediction phase in the existing studies. However, the deep learning models, in particular, Long Short-Term Memory (LSTM) networks, are well-known models for modelling sequential climate data.

Climate information can also vary greatly with time, having seasonal cycles and long-term variations, which can be difficult to capture in machine learning. Alternative networks are based on learning temporal relations in time-series data; thus, they are applicable for climate prediction. Some hybrid and alternative models have been investigated, but there are always reports of good performance in temporal pattern capturing using LSTM models.

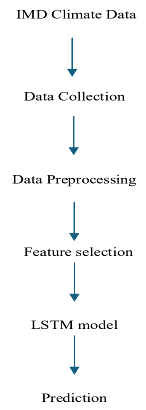

Fig. 3. Proposed Methodology Framework for Climate-Based Prediction

c. Framework Interpretation

The conceptual framework (Fig. 3) is a structured process to connect climate data processing to predictive modelling and agricultural decision making. Combines pre-processing, feature extraction and time-series modeling in a single pipeline. It is not a specific system in place, but a general guideline on the basis of previous studies. It also highlights the need to further the existing trend of multi-source data integration as an important research avenue.

VI. TIME-SERIES MODELING FOR CLIMATE PREDICTION

The climate data is naturally sequential and has a high temporal correlation, such as in the form of phenology, seasonal rainfall patterns and temperature changes, as well as long-term trends in climate. This would be difficult to model through the traditional machine learning models with the assumption of independent observations.

This, however, requires that the techniques used to model the time series need to be employed to overcome this limitation. From the subcategories of recurrent neural networks, the use of Long Short-Term Memory (LSTM) networks seems best suited since they can retain knowledge for extended periods of time within sequences of temporal data and can also be trained to extract complex temporal patterns. LSTM is a technique that regulates the flow of information by using gating units and memory cells such as the Input Gate, Forgotten Gate, Output Gate.

LSTM models are powerful for climate-based predictions as they can capture a seasonal pattern and long-term prediction from multiple datasets, thus improving prediction accuracy. They cannot be employed in the context of multi-source climate data systems, however, and that opens a myriad of research opportunities.

VII. FUTURE DIRECTIONS

There are several areas in need of further research into the future:

- Integrating several data sources like IMD, IITM, NIO, etc.
- Using LSTM for statistical analysis of time series data.
- To construct computer-based pipelines of climate predictions.
- To build automatic lines for climate prediction.

- Applying real-time data to a more dynamic prediction – just one of these applications

These advances can make a substantial contribution to the ability of climate prediction systems to be accurate and efficient.

VIII. CONCLUSION

This is a review article on different approaches to climate prediction in agriculture systems in order to summarize and critically examine the studies published. The review reveals that the field of climate prediction science has been using AI techniques, such as machine learning and deep learning models, to enhance the ability to predict. The review shows that the use of machine learning and deep learning models in climate prediction research is on the rise, and this has helped greatly in increasing the accuracy of climate predictions. LSTM models appear to be a suitable approach for dealing with time-series related to climate, because of their ability to model temporal dependence. Some common issues identified during the review are the absence of temporal modelling in the traditional solutions, absence of data integration, complexity of the pre-processing, and absence of scalable and deployment-oriented frameworks. In this study, IMD data is taken as primary structured data and other data that are not covered by IMDS are discussed as IITM and NIO. To organize these results, a conceptual framework has been provided, with data preprocessing and feature extraction as highlighted tasks, and modeling of time-series using LSTM as the key process. The framework could be expanded in the future by incorporating further data sources of climate and providing predictive services for agriculture in the face of climate change.

REFERENCES

[1] S. Ghosh, M. Das, and A. Chakraborty, "Crop Yield Prediction Using Supervised Machine Learning Techniques: A Case Study from India," in Proc. ICCCIS, pp. 694–698, 2020.

[2] V. Kumar and S. Bandyopadhyay, "An End-to-End Framework for Real-Time Data Preprocessing and Machine Learning Model Deployment in Smart Farming," Procedia Computer Science, vol. 189, pp. 275–282, 2021.

[3] P. Sharma, A. K. Meena, and R. K. Tripathi, "Cloud-Based Smart Irrigation System Using Machine Learning and Automation," Journal of Ambient Intelligence and Humanized Computing, vol. 12, no. 8, pp. 8679–8691, 2021.

[4] R. Rajan, S. Kaur, and A. Kumar, "MLOps Pipeline for Agricultural Applications: A Modular Approach," International Journal of Computer and Information Engineering, vol. 14, no. 10, pp. 418–425, 2020.

[5] N. Jain and S. Choudhary, "Automated Data Preprocessing Techniques for Real-Time Crop Monitoring Using AI," International Journal of Artificial Intelligence in Agriculture, vol. 3, no. 1, pp. 29–36, 2022.

[6] L. Zhang, T. Yu, and Y. Chen, "Towards Explainable AI in Agriculture: A Survey on Model Interpretability Techniques," IEEE Access, vol. 10, pp. 45821–45835, 2022.

[7] B. Nguyen, H. Tran, and Q. Pham, "IoT-Based Framework for Real-Time Crop Monitoring Using Deep Learning," Computers and Electronics in Agriculture, vol. 193, 106624, 2022.

[8] M. Tiwari and D. Verma, "A Comparative Analysis of Machine Learning Algorithms for Precision Agriculture," International Journal of Advanced Computer Science and Applications, vol. 13, no. 6, pp. 180–186, 2022.

[9] S. Banerjee and A. Sinha, "Explainable ML Models in Agricultural Decision Support Systems," IEEE Access, vol. 10, pp. 122545–122557, 2022.

[10] L. Wang and X. Liu, "A Novel Crop Recommendation System Using Deep Learning Techniques," Journal of Agricultural Informatics, vol. 14, no. 1, pp. 25–36, 2023.

[11] R. Patel and M. Sharma, "XAI-Based Crop Recommendation System Using Explainable Artificial Intelligence," in Proc. Int. Conf. Computational Intelligence and Smart Agriculture, pp. 148–155, 2023.

[12] A. Khan and S. Roy, "Advanced Ensemble Learning Methods in Crop Recommendation Systems," in Proc. IEEE Conf. Intelligent Systems in Agriculture, pp. 93–99, 2023.

[13] T. Singh and R. Gupta, "Synthetic Data Generation in Smart Farming Using GAN and VAE," Artificial Intelligence in Agriculture, vol. 5, pp. 78–90, 2023.

[14] K. Rana, R. Kumar, and M. Singh, "AI-Based Crop Disease Diagnosis System," International Journal of Artificial Intelligence in Agriculture, vol. 4, no. 2, pp. 56–64, 2023.

[15] R. Chatterjee and V. Patel, "Smart Farming Using ML and Explainable AI," in Proc. IEEE AgriTech Digital Farming Conf., pp. 201–208, 2023.

[16] D. R. Sharma and S. Jain, "MLOps for Sustainable Agriculture," Journal of Intelligent & Fuzzy Systems, vol. 45, no. 1, pp. 591–601, 2023.

[17] M. Hossain, T. Alam, and S. Uddin, "Hybrid ML Approach for Yield and Climate-Based Prediction," Information Processing in Agriculture, vol. 10, no. 2, pp. 237–246, 2023.

[18] A. Bhatt, R. Das, and S. Mehta, "Deep Learning in Smart Farming: From Data Collection to Decision-Making," ACM Computing Surveys, vol. 56, no. 3, pp. 45:1–45:35, 2024.

[19] M. A. Jabed and M. A. Murad, "Crop Yield Prediction in Agriculture: A Comprehensive Review of Machine Learning and Deep Learning Approaches," Heliyon, vol. 10, no. 24, 2024.

[20] A. C. Cob-Parro, Y. Lalangui, and R. Lazcano, "Fostering Agricultural Transformation through AI: An Open-Source AI Architecture Exploiting the MLOps Paradigm," Agronomy, vol. 14, 2024.

[21] R. N. V. Jagan Mohan et al., "Next-Gen Agriculture: Integrating AI and Explainable AI for Precision Crop Yield Prediction," Frontiers in Plant Science, 2025.

[22] C. R. Screpnik, E. Zamudio, and L. I. Gimenez, "Artificial Intelligence in Agriculture: A Systematic Review of Crop Yield Prediction and Optimization," IEEE Access, 2025.

[23] S. Gupta, S. Mohanty, and D. K. Behera, "AI-Based Yield Prediction: A Thorough Review," Indian Journal of Science and Technology, vol. 18, no. 10, pp. 822–838, 2025.

[24] K. Jhajharia, N. V. Sharma, and P. Mathur, "A Machine Learning Model for Crop Yield Prediction Using Remote Sensing Data," Int. Research Journal of Multidisciplinary Scope, vol. 6, no. 2, pp. 577–590, 2025.

[25] R. K. Dhanaraj et al., "On-Device AI for Climate-Resilient Farming with Intelligent Crop Yield Prediction," Scientific Reports, vol. 15, 2025.

[26] A. K. Pal, B. D. K. Patro, and S. Chaube, "Design and Implementation of a Deep Learning Framework for Automated Crop Classification and Health Diagnosis," Scientific Reports, 2026.