Climate-Driven Crop Yield Forecasting: A Review of Machine Learning, Deep Learning, and Remote Sensing Approaches

Kapil Keshao Wankhade[1], Ganesh Khekare[2]

[1] Lincoln University College, 47301, Petaling Jaya, Selangor Darul Ehsan, Malaysia; [2] School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, 632014, Tamil Nadu, India

Email ID [1]**kaps.wankhade@gmail.com; [2]khekare.123@gmail.com**

Abstract: Effective crop yield prediction is also crucial in food security, agricultural management, and reducing the effects of climate variability. Over the recent years, data analytics on climate, coupled with artificial intelligence tools, have been enhanced to a great extent prediction abilities of crop yield models. The paper is a review of the position of crop yield prediction methods that utilize climate variables, deep learning, machine learning, and remote sensing methods. The review conducts a systematic study of both classical statistical approaches and current regression-based machine learning models, as well as classification models, such as neural networks, gradient boosting, support vector machines, and random forests. More so, long short-term memory networks and convolutional neural networks are covered under the deep learning algorithm. The research paper identifies the frequently employed climate variables of temperature, rainfall, soil moisture, humidity, and the multi-source data available. The most important issues, such as data heterogeneity, time uncertainty, climate extremes, and generalization of models across regions, are highly discussed. Lastly, the paper also provides future research directions with a focus on multi-modal data fusion, explainable AI, and scalable forecasting frameworks in order to make them robust and relevant in the real world.

Keywords: Machine learning; Remote sensing; Deep learning; Climate variables; Crop yield prediction

Introduction

In the world, one of the most climate sensitive sectors is agriculture, which directly affects food security, rural livelihood, and stability in the global economy. Environmental conditions are very strict on crop yield variability, and any insignificant changes in climatic parameters can lead to significant losses. As climatic variability, extreme weather conditions, and long-term climatic changes continue to increase, precise knowledge and prediction of crop yield is now of great importance to the sustainable agricultural planning and reduction of risks. This has led to the emergence of more scientific studies in determining the interactions between climatic factors and crop productivity to improve resilience to yield and decision making.

Some of the most dominant environmental stressors that influence crop growth and productivity include salinity and drought. Salinity stress disturbs physiological functions, including the uptake of water, ions, and photosynthesis, which eventually causes low crop yield. Zorb et al. [1] were able to systematically review the effects of salinity on crop yield, and emphasized the salt effects on osmotic stress and ion toxicity that largely inhibit plant growth in different crop species. On the same note, drought stress has been cited as a principal constraint to agricultural production, especially in rain-fed areas. Dietz et al. [2] established that long-term water scarcity alters metabolic processes and reduces biomass accumulation, leading to significant yield loss. In addition to the abiotic stresses, the crop yield is a result of a complicated relationship between economic, site-specific, and climatic factors. It was highlighted by Cabas et al. [3] that the yield response cannot be described as the result of the climate factors only, but also includes the soil properties, the management, and the economic factors. The results of their research revealed that the precipitation and temperature variability interact with the local site factors to influence the outcomes of crop yields, which has led to the importance of combined modeling strategies that can be used to observe the presence of both environmental and contextual factors.

The recent research studies have also indicated that the relationships between climate and crop yields are dynamic and change with time as a result of climate change and adaptation mechanisms. Feng et al. [4] compared long-term climate records and yield records and found that the changing temperature and rainfall pattern has changed the risk profile of yield decrease in various regions. These changing relationships make the traditional forecasting models less certain and require some data-driven methods that are adaptive and can capture non-linear and dynamic interactions. Najafi et al. [5] examined the behavior of yields across the globe and concluded that, inasmuch as there have been improvements in technology, which have led to an increase in yields, the variability in climate is still limiting yield production, especially in the developing world.

The organization of the paper is that the next section, related work, further studies examine deep learning and machine learning methods in combination with remote sensing methods as a way to solve the traditional model's limitations and enhance predictive accuracy due to the uncertainty of climate conditions. This is followed by the last section, which is the Conclusion, which sums up the paper.

Related work

Crop yield forecasting is a crucial tool to maintain food security across the globe, sustainable agricultural growth, and good management of the available resources as climate variability increases. Farm output is a naturally sensitive area to climatic conditions, and changes in temperature, rainfall, humidity, and weather extremes have a direct effect on the growth, development, and ultimate yielding of crops.

A. Climate-Based Crop Yield Forecasting

Crop yield prediction based on climatic conditions has become a considerable issue in light of the rising level of climate variability and the direct relationship between climatic variability and agricultural output. A framework of predictive analytics in crop yield, based on machine learning and climate-driven predictive analytics, was proposed by Jaikrishna et al. [6]. Historical climate variables like temperature, rainfall, and humidity were used in their study to forecast the trends in crop yields. The biggest advantage of this method is that it is capable of capturing non-linear interactions between climatic factors and yield results. The performance of the model is, however, very sensitive to the quality of data and time consistency. Dhanke et al. [7] proposed an AI-based precision irrigation system to maximize the crop yield under climate-induced limitations. Their strategy involved a combination of climate data and smart irrigation decision-making in order to make things more sustainable. The merit of this work is that it is a holistic combination of climate analytics and agricultural management practices.

Niedballa et al. [8] have given an extensive discussion on methods of prediction and estimation of agricultural production under varying climatic conditions. The paper has examined various forecasting models and how predictive systems should be modified to match changing climatic patterns.

B. Traditional and Statistical Models

The basis of crop yield prediction studies is traditional and statistical models, which have been extensively used to analyse relationships between climate and yield since they are easily interpretable and simple.

Lobell et al. [9] examined how statistical models could be used to forecast climate change on crop yields. They used regression-related methods in their work to measure the effects that temperature and precipitation have on the variability of the yield. Shi et al. [10] discussed an extensive survey of statistical models that have been employed to determine the contributions of climate to crop yields. Linear regression, panel data model, and econometric techniques were noted to be commonly used in the study. In the case of wheat in Central Punjab, Gill et al. [11] discussed a statistical model based on the weather used to predict the yield. The model was able to include seasonal weather indicators to predict yield outcomes. Its greatest advantage is the fact that it was region-specifically calibrated, and it enhanced prediction accuracy in its region.

C. Machine Learning Models

Machine learning (ML) methods have greatly improved crop yield estimation through the modeling of non-linear and complex relationships between climate factors and crop production.

Veenadhari et al. [12] paid attention to regression modeling based on the climate, in which temperature, rainfall, and humidity may be mapped to provide values. Many datasets can be combined with the help of high computation and merged into a single dataset from which all the important features can be extracted. Jeong et al. [13] expanded this idea to the global and regional levels with the help of Random Forest regression and emphasized its resilience to noisy climate data as well as its capabilities to deal with the high-dimensional input.

Table 1 shows the general workflow followed in machine learning-based crop yield prediction. A single climate dataset can undergo either regression (continuous yield forecasting) or be transformed into a set of classification labels (yield categories), which proves that ML models are versatile for a variety of forecasting goals.

Dang et al. [14] compared regression and classification properties between ML and deep learning models on climatic data and remote sensors features, and found the performance of the ensemble and neural models to be higher than that of single learners in climate variability.

Table 1. General Machine Learning Pipeline for Climate-Based Crop Yield Forecasting

Step	Description (Regression & Classification)
Dataset Loading	Climate variables such as temperature, rainfall, humidity, soil moisture, and historical crop yield data are collected
Data Preprocessing	Missing value handling, normalization or standardization, and removal of noisy or outlier data
Feature Selection	Selection of relevant climate indicators and derived temporal or statistical features
Label Formation	Continuous yield values for regression or discretized yield classes for classification

Model Training	Training of ML models, including Artificial Neural Networks, Logistic Regression, KNN, SVM, and Random Forest
Model Evaluation	Evaluation using RMSE and MAE for regression, and F1-score, recall, precision, and accuracy for classification
Prediction	Forecasting future crop yield values or yield categories using unseen climate data

D. Deep Learning Models

Deep learning (DL) models have become highly effective in predicting crop yields because they can discover hierarchical representations of challenging and high-dimensional climate and remote sensing data on their own.

Paudel et al. [15] aimed at the interpretability of deep learning models, which is a significant problem in the field of agricultural AI due to the constraints of black-box predictions in the real world. The techniques of explainability they suggest were created to learn the impact of climate characteristics on yield predictions. Conversely, Sabo et al. [16] tested the deep learning designs in small data scenarios and found that more complex models do not necessarily indicate high accuracy. Though not as effective as other types of models, DL models can be trained using large datasets and need thorough regularization to prevent overfitting.

Khaki et al. [17] proposed a CNN-RNN hybrid architecture, where spatial features of climate or vegetation data are extracted with Convolutional Neural Networks and time-dependent yield dynamics are learn by Recurrent Neural networks.

Conclusions

This review systematically reviewed the studies on climate-driven crop yield forecasting, as they undergo an evolution from traditional statistical-based forecasting models to modern-day remote sensing-based forecasting models, deep learning, and machine learning. The results of the analysis indicate that machine learning algorithms like the SVM, Random Forest, and ensemble methods are effective to learn nonlinear relations between climate variables and crop yield, whereas deep learning models such as CNNs, RNNs, and attention-based models can learn spatio-temporal relationships with multisource climate and satellite data. Although these developments have been made, issues concerning data heterogeneity, propagation of uncertainty, interpretability of the models, and cross-regional generalization still constrain large-scale implementation and confidence in the decisions made by the decision maker.

The next step in future research must be to investigate uncertainty-conscious and explainable hybrid systems that incorporate climate forecasts, remote sensing, and agronomic knowledge to make them more robust and transferable.

References

1. C. Z¨orb, C.-M. Geilfus, and K.-J. Dietz, “Salinity and crop yield,” Plant Biology, vol. 21, pp. 31–38, 2019.
2. K.-J. Dietz, C. Z¨orb, and C.-M. Geilfus, “Drought and crop yield,” Plant Biology, vol. 23, no. 6, pp. 881–893, 2021.

3. J. Cabas, A. Weersink, and E. Olale, “Crop yield response to economic, site and climatic variables,” Climatic Change, vol. 101, no. 3, pp. 599–616, 2010.
4. S. Feng, Z. Hao, X. Zhang, and F. Hao, “Changes in climate–crop yield relationships affect risks of crop yield reduction,” Agricultural and Forest Meteorology, vol. 304, Art. no. 108401, 2021.
5. E. Najafi, N. Devineni, R. M. Khanbilvardi, and F. Kogan, “Understanding the changes in global crop yields through changes in climate and technology,” Earth’s Future, vol. 6, no. 3, pp. 410–427, 2018.
6. M. Jaikrishna, S. H. Kumar, and G. Revathy, “Predictive analytics for climate-driven crop yield forecasting using machine learning,” in Proc. 8th Int. Conf. Circuit, Power & Computing Technologies (ICCPCT), IEEE, 2025, pp. 757–762.
7. J. A. Dhanke et al., “Climate-based AI-powered precision irrigation: Sustainably smart agriculture frameworks for maximum crop yields,” Remote Sensing in Earth Systems Sciences, vol. 8, no. 1, pp. 161–172, 2025.
8. G. Niedbała, M. Piekutowska, T. Wojciechowski, and M. Niazian, “Predictions and estimations in agricultural production under a changing climate,” Agronomy, vol. 14, no. 2, Art. no. 253, 2024.
9. D. B. Lobell and M. B. Burke, “On the use of statistical models to predict crop yield responses to climate change,” Agricultural and Forest Meteorology, vol. 150, no. 11, pp. 1443–1452, 2010.
10. W. Shi, F. Tao, and Z. Zhang, “A review on statistical models for identifying climate contributions to crop yields,” Journal of Geographical Sciences, vol. 23, no. 3, pp. 567–576, 2013.
11. K. K. Gill et al., “Wheat yield prediction using weather based statistical model in Central Punjab,” Journal of Agricultural Physics, vol. 15, no. 2, pp. 157–162, 2015.
12. S. Veenadhari, B. Misra, and C. D. Singh, “Machine learning approach for forecasting crop yield based on climatic parameters,” in Proc. Int. Conf. Computer Communication and Informatics, IEEE, 2014, pp. 1–5.
13. J. H. Jeong et al., “Random forests for global and regional crop yield predictions,” PLOS ONE, vol. 11, no. 6, Art. no. e0156571, 2016.
14. C. Dang, Y. Liu, H. Yue, J. Qian, and R. Zhu, “Autumn crop yield prediction using data-driven approaches,” Canadian Journal of Remote Sensing, vol. 47, no. 2, pp. 162–181, 2021.
15. D. Paudel et al., “Interpretability of deep learning models for crop yield forecasting,” Computers and Electronics in Agriculture, vol. 206, Art. no. 107663, 2023.
16. F. Sabo, M. Meroni, F. Waldner, and F. Rembold, “Is deeper always better? Evaluating deep learning models for yield forecasting with small data,” Environmental Monitoring and Assessment, vol. 195, no. 10, Art. no. 1153, 2023.
17. S. Khaki, L. Wang, and S. V. Archontoulis, “A CNN–RNN framework for crop yield prediction,” Frontiers in Plant Science, vol. 10, Art. no. 1750, 2020.