ASSESSMENT OF CROP YEILD PREDICTION IN INDIAN SCENARIO: **A REVIEW**

**Deshmukh Nilay1, Deshpande Santosh2, Rokade Prathmesh3, Ghatage Shreyash4, Dr. Krantikumar. V. Mhetre5**

**1, 2, 3, 4 Students, Civil Engineering, APCOER, Pune, Maharashtra, India.**

**5Assistant Professor, Civil Engineering, APCOER, Pune, Maharashtra, India.**

**ABSTRACT**

Studying the literature of differential modelling techniques used to develop models from various crop data makes a significant contribution to the agriculture industry. The increasing concerns about climate change and environmental sustainability have directed significant attention towards agriculture. This study conducts a comprehensive literature review to analyze current research trends, methodologies, new research techniques and findings related to crop yield predictions. It aims to identify the existing literature of agriculture industry and gap identify to guide future research efforts.

**Keywords:** Crop yield prediction, Agriculture, Environmental impact, Climate

1. **INTRODUCTION**

Agriculture is one of the most important activities that produces crop and food that is crucial for the sustenance of a human being. Agriculture gave birth to civilization in India. India is an agrarian country, and its economy based upon crop productivity. Thus, agriculture is the backbone of all business in India. The entire economy also depends on the produce from harvesting annually. In the present day, agricultural products and crops are not only used for local demand, but globalization has allowed us to export produce to other countries and import from other countries. India is an agricultural nation and depends a lot on its agricultural activities.

Today biggest challenges of farmers are determining which crops suitable for the soil, what type of fertilizers to be use, how to deal with unexpected climatic changes and how to make crop production profitable by using minimum inputs. Since crop yield prediction is one of the challenging problems in precision agriculture and many models have been proposed and validated so far. This prediction will help farmers from our study area to choose whether the crop is suitable for specific season and available water and soil characteristics. Our Aim is to predict crop yield based on the water uses, soil characteristics and seasonal requirement. Along with it we also aim to predict pH value of various soil types in study area.

Crop yield prediction is the main task that is significant in our study area. Traditionally, farmers predict the crop production and yield by considering the rainfall trends and the number of crops sown, but the change in climate as well as the fundamental complexity of factors that influence the crop production have made the prediction more difficult and less accurate.

To overcome these limitations, there are other ways by which prediction of crop yield can be done. In our study area, we are focused on analysis, such as water uses, soil characteristics, soil pH value, season wise crop yield prediction.

1. **METHODOLOGY**

This methodology is based on crop yield prediction using literature of crop yield prediction from different locations.

**Literature Review**

**Conclusion**

**Identify Literature Gap**

**Topic Finalization**

**Figure 1** Methodology

1. **LITERATURE REVIEW**

Wang et al. (2024) reviewed precision agriculture's use of remote sensing and machine learning to address challenges like resource scarcity and climate change in agriculture. Hyperspectral and UAV remote sensing are widely adopted, with Support Vector Machines and Random Forests leading machine learning applications. Key challenges include data quality and model interpretation, with future directions focusing on advancing agricultural tech and fostering international collaboration, [12].

Wang et al. (2024) developed high-resolution datasets for maize and wheat yields and crop water productivity (CWP) in China to improve agricultural predictions. Using remotely sensed indicators and a random forest model, they created 1-km resolution data, showing MOD16 as a reliable tool for evapotranspiration measurement. The model achieved RMSEs of 26.81% for maize and 21.80% for wheat locally, with improved accuracy regionally. These findings aid understanding of crop yield and water productivity patterns, supporting better agricultural strategies and food security, [13].

Uppugunduri et al. (2024) reviewed machine learning's role in crop yield prediction, a complex process influenced by factors like soil, temperature, and rainfall. Analyzing 660 papers, they selected 50 that demonstrated effective methods and accuracy. Their review identifies key features, methods, and geographic contexts for yield prediction, highlighting research gaps and providing insights to guide future improvements in predictive models, [9].

Miranda et al. (2024) tackled crop yield prediction using Remote Sensing (RS) with a novel Multi-view Gated Fusion (MVGF) model. Integrating data from sources like Sentinel-2 imagery, weather, and soil properties, the model enhanced prediction accuracy for crops in Argentina, Uruguay, and Germany. With a 10 m resolution, the MVGF model achieved R² scores of 0.68 for sub-field predictions in Argentina and about 0.80 across countries, showing significant potential for precision agriculture, [7].

Mena et al. (2024) explored cropland analysis challenges with Multi-View Learning (MVL) models, focusing on combining diverse data types. Testing five fusion strategies and five temporal encoders on the Crop Harvest dataset, they found that a tailored approach, rather than one-size-fits-all, better handles data diversity, improving crop classification. Their insights guide optimal MVL configurations for advanced agricultural analysis, [5].

Jhajharia et al. (2023) developed an interaction regression model for crop yield prediction, crucial for food security. This model, applied to corn and soybean yields in Illinois, Indiana, and Iowa, achieved a low relative RMSE of 8% and provided insights into how environmental and farming factors impact yields. It helps agronomists understand yield influences from weather, soil, and management, offering accurate predictions and clear guidance for agricultural decision-making, [4].

Sutha et al. (2023) studied climate change's effect on water needs for Rajasthan’s monsoon and winter crops. Using the Penman-Monteith equation, they estimated crop water requirements through 2100, finding that warming by 4°C could increase water needs by 12.9–19.9% across crops like millet, bean, and mustard. This rise may also shorten growth periods for longer crops and significantly reduce rainfall-sufficient areas for crops, with a severe impact on groundwater-reliant winter crops, [11].

Villalobos and Morales (2023) conducted a systematic review of deep learning in crop yield prediction, selecting 44 studies out of 456 for their quality. They found Convolutional Neural Networks (CNNs) to be the most effective in terms of RMSE but noted challenges like overfitting due to limited large datasets. The review calls for more research to improve deep learning applications in this field, [8].

Halder et al. (2023) focused on improving crop yield predictions in India using machine learning, especially for climate-sensitive crops like rice and cotton. They evaluated various ML models and found the Extra Trees Regressor performed best with an R-squared of 0.9615, followed by Random Forest and LGBM, highlighting the value of tree-based models for agricultural predictions, [3].

 Mena et al. (2023) explored the complexities of crop classification with remote sensing (RS) data, noting challenges from varying data resolutions and noise. The study assessed fusion techniques on the Crop Harvest dataset, finding that no single fusion method worked best across all contexts. Instead, performance varied by region, highlighting the need for context-specific fusion strategies in multi-view learning. They proposed a framework to help researchers choose optimal fusion methods for different datasets, advancing RS-based crop classification accuracy, [6].

Oikonomidis et al. (2022) explored the use of machine learning (ML) in crop yield prediction, emphasizing the effectiveness of Random Forests over other algorithms like artificial neural networks, with an RMSE of 35-38%. They found that data organization plays a crucial role in improving prediction reliability, and models based on ordered data perform better than those with random partitioning. Despite this, even top ML models showed only modest improvements over basic yield predictions, raising questions about the cost-effectiveness of complex data collection. The study advocates for comparing ML methods with simpler approaches and using validated crop models before data collection, [10].

Burgess et al. (2022) focused on maize yield prediction, considering the complex interactions between genotype and environmental factors. Their winning model, a deep neural network (DNN), achieved an RMSE of 12% using predicted weather data. The model effectively used feature selection to reduce input complexity while maintaining accuracy, outperforming traditional methods like Lasso and regression trees. The study highlighted that environmental factor had a greater impact on yield than genetic traits, underscoring the importance of incorporating environmental variability in forecasting and the potential of deep learning in improving agricultural predictions, [1].

Cheng et al. (2022) highlights the critical role of accurate crop yield prediction in agricultural science for informing national economic policies and decision-making. While traditional machine learning methods have been used, they often lack precision, leading to discrepancies between predicted and actual yields. The paper explores deep learning advancements, which show promise in improving prediction accuracy. It provides a comprehensive review of existing models for different crops, discussing their strengths and weaknesses. The study also suggests future research directions, advocating for the integration of deep learning to enhance yield forecasting, ultimately supporting agricultural sustainability and food security, [2].

1. **LITERATURE GAP**

We went through 13 research papers and noticed that researchers have used a range of modelling techniques along with inputs like remote sensing images, soil properties, rainfall data, and climate patterns to get their results. But interestingly, none of them seemed to take a more holistic approach—like predicting crop yield based on a mix of factors such as soil pH, organic matter, nutrients like nitrogen, phosphorus, and potassium, along with crop prices, fertilizer usage, overall production, and water consumption. This shows there’s still room to explore how combining all these elements could lead to more accurate and insightful yield predictions.

1. **RESULTS AND DISCUSSION**

Below table 1 shows the literature study conducted for the research paper divided in study area, author names, year of study and reference no of their research.

**Table 4.1** Study the literature

|  |  |  |  |
| --- | --- | --- | --- |
| **Author Name** | **Year** | **Country** | **Reference No** |
| Wang et al. | 2024 | Global | 12 |
| Wang et al. | 2024 | China | 13 |
| Uppugunduri et al. | 2024 | India | 9 |
| Miranda et al. | 2014 | Argentina, Uruguay, Germany | 7 |
| Mena et al. | 2024 | Global | 5 |
| Jhajharia et al. | 2023 | India | 4 |
| Sutha et al. | 2023 | India | 11 |
| Villalobos and Morales | 2023 | United States, India, China | 8 |
| Halder et al. | 2023 | Global | 3 |
| Mena et al. | 2023 | Global | 6 |
| Oikonomidis et al. | 2022 | Global | 10 |
| Burgess et al. | 2022 | Global | 1 |
| Cheng et al. | 2022 | China | 2 |

1. **CONCLUSION**

Study the literature of crop yield prediction makes a significant contribution to the agricultural industry. The increasing concerns climate change and Impact on climate. literature gap is fulfilled by differential factors can be use for the future study. By Combining various factors there is still room to explore for the future studies.

**ACKNOWLEDGEMENTS**

We thank Dr. Krantikumar V. Mhetre from Anantrao Pawar college of engineering and research, Pune who provides insight and expertise that greatly assisted the work. Secondly, we would also like to thank Dr. R. R. Sorate Head of the Civil Engineering department and all staff members and our all colleagues, for their valuable suggestions and support. We would like to thank Dr. S. B. Thakare sir, Principal of our institute APCOER Pune.

**REFERENCES**

* 1. Burgess, A. J., Masclaux‐Daubresse, C., Strittmatter, G., Weber, A. P., Taylor, S. H., Harbinson, J., Yin, X., Long, S., Paul, M. J., Westhoff, P., Loreto, F., Ceriotti, A., Saltenis, V. L., Pribil, M., Nacry, P., Scharff, L. B., Jensen, P. E., Muller, B., Cohan, J., Foulkes, J., Rogowsky, P., Debaeke, P., Meyer, C., Nelissen, H., Inzé, D., Klein Lankhorst, R., Parry, M. A., Murchie, E. H., and Baekelandt, A. (2022). “Improving crop yield potential: Underlying biological processes and future prospects.” Food and Energy Security, 12(1).
	2. Cheng, M., Jiao, X., Shi, L., Penuelas, J., Kumar, L., Nie, C., Wu, T., Liu, K., Wu, W., and Jin, X. (2022). “High-resolution crop yield and water productivity dataset generated using random forest and remote sensing.” Scientific Data, 9(1).
	3. Halder, M., Datta, A., Siam, M. K. H., Mahmud, S., Sarkar, Md. S., and Rana, Md. M. (2023). “A Systematic Review on Crop Yield Prediction Using Machine Learning.” Research Gate, 177.
	4. Jhajharia, K., Mathur, P., Jain, S., and Nijhawan, S. (2023). “Crop yield prediction using machine learning and Deep Learning Techniques.” Procedia Computer Science, 218, 406–417.
	5. Mena, F., Arenas, D., and Dengel, A. (2024). “IN THE SEARCH FOR OPTIMAL MULTI-VIEW LEARNING MODELS FOR CROP CLASSIFICATION WITH GLOBAL REMOTE SENSING DATA.” IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2024, with pages 4307–4311.
	6. Mena, F., Arenas, D., Nuske, M., and Dengel, A. (2023). “A comparative assessment of Multi-View Fusion Learning for Crop Classification.” IGARSS 2023 - 2023 IEEE International Geoscience and Remote Sensing Symposium, 5631–5634.
	7. Miranda, M., Pathak, D., Nuske, M., Dengel, A., Mena, F., Najjar, H., Sanchez, C., Siddamsetty, J., Charfuelan, M., Arenas, D., Vollmer, M., Helber, P., Bischke, B., and Habelitz, P. (2024). “ADAPTIVE FUSION OF MULTI-VIEW REMOTE SENSING DATA FOR OPTIMAL SUB-FIELD CROP YIELD PREDICTION.” IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, 4307–4311.
	8. Morales, A., and Villalobos, F. J. (2023). “Using machine learning for crop yield prediction in the past or the future.” Frontiers in Plant Science, 14.
	9. Nikhil, U. V., Pandiyan, A. M., Raja, S. P., and Stamenkovic, Z. (2024). “Machine learning-based crop yield prediction in South India: Performance analysis of various models.” Computers, 13(6), 137.
	10. Ofori-Ampofo, S., Pelletier, C., and Lang, S. (2021). “Crop type mapping from optical and radar time series using attention-based deep learning.” Remote Sensing, 13(22), 4668.
	11. Sutha, K., Indumathi, N., and Uma Shankari, S. (2023). “Recommending and predicting crop yield using Smart Machine Learning Algorithm (SMLA).” Current Agriculture Research Journal, 11(2), 686–694.
	12. Wang, J., Wang, Y., Li, G., and Qi, Z. (2024). “Integration of Remote Sensing and machine learning for Precision Agriculture: A comprehensive perspective on applications.” Agronomy, 14(9), 1975.
	13. Wang, Y., Zhang, Q., Yu, F., Zhang, N., Zhang, X., Li, Y., Wang, M., and Zhang, J. (2024b). “Progress in research on Deep Learning-based crop yield prediction.” Agronomy, 14(10), 2264.