Elinvo (Electronics, Informatics, and Vocational Education) ISSN 2580-6424 (printed) | ISSN 2477-2399 (online)
Vol 11, No. 1, 2026, pp. 34 - 42

Data-Driven Assessment of Rice Yield Gaps in Rainfed Agriculture Using Predictive Modeling and Cluster Analysis

Heru Ismanto *1, Lilik Sumaryanti 2, Daud Andang Pasalli 3

123University Musamus, Merauke, Indonesia
1 heru@unmus.ac.id*; 2 lilik@unmus.ac.id; 3 daud@unmus.ac.id
* corresponding author

Article Info

Article history:
Received Februari 09, 2026
Revised June 09, 2026
Accepted June 22, 2026
Available Online June 23, 2026

Keywords:
Yield Gap; Rainfed Rice;
Machine Learning;
Efficiency Clustering;
Agricultural Policy;

Abstract

Optimizing agricultural productivity in marginal areas of Merauke Regency, South Papua Province, Indonesia, faces significant challenges due to a high yield gap and low input efficiency. This study proposes an innovative machine learning-based approach to evaluate and map the performance of upland rice farmer groups by using PFPL (Prospective Farmer Prospective Location) data, which only has previously been used administratively. By integrating a predictive model (Random Forest Regressor), success classification, and K-Means Clustering, this study builds an adaptive and replicable analytical framework to support Data-Driven agricultural decision-making. The analyzed dataset includes 30 farmer groups which containing technical information such as land area, seed use, pesticide use, and herbicide use, as well as actual and targeted yields. The feature engineering process yielded the input efficiency ratio as the primary variable. The Random Forest regression model achieved a near-perfect fit on the available dataset (R^2 = 0.95; RMSE = 0.41). However, given the limited sample size (30 farmer groups), the result should be interpreted cautiously and regarded as exploratory rather than conclusive. Cluster analysis revealed two segments: a high-input but inefficient group and an efficient group with very high yields. These results highlight that input quantity does not guarantee productivity without efficient use. This study not only expands the literature on agricultural intelligence but also offers a practical approach for policymakers to design efficiency-based interventions, incentives, and training. This approach is also relevant for accelerating digital transformation and food security in underdeveloped regions.

Corresponding Author:
Email: heru@unmus.ac.id

INTRODUCTION

Agriculture is a crucial sector in Indonesia because it absorbs labor, directly and indirectly impacting to the economy of Indonesia [1]. One important agricultural commodity in Indonesia is rice. Rice is a strategic food commodity in Indonesia that supports national food security and the welfare of farmers, especially in rural areas. One of the main challenges in developing rice farming is the yield gap between actual farmer productivity and the potential yield that can be achieved through optimal cultivation practices [2]. This problem is increasingly complex in dryland and rain-fed areas, such as Papua Province in general [3], [4] and South Papua, Indonesia in particular, which have marginal agro-ecological conditions and are often not served by irrigation systems, leading to declining agricultural production [5]. Despite various government intervention programs, the yield gap remains significant, and conventional approaches to evaluating farmer group performance tend to be static, non-adaptive, and not comprehensively Data-Driven [6].

The main problem which was highlighted in this study is the suboptimal utilization of PFPL (Prospective Farmer Prospective Location) data to analyze input efficiency patterns and yield performance of farmer groups in the rainfed area of Merauke Regency, South Papua. This data, which includes information on land area, seed varieties, pest types, and actual yields and productivity targets, has been used primarily for administrative purposes, rather than predictive analytics. This case gives effects to some aspects such as the decisions about aid distribution, technical training, and intervention programs not being fully based on real needs or local potential [7]. Therefore, a new approach is needed which can translate this data into meaningful insights by using machine learning and spatial analytics technology.

The previous studies have demonstrated the potential of machine learning in predicting crop yields [8], [9], [10], [11], grouping regions based on agroecological similarities [12], [13], predicting rice yields [14], [15], [16], [17], [18], and identifying agricultural input efficiency [19], [20], [21]. However, most of these studies have focused on irrigated agricultural systems in developed countries, and there are only a quite few studies have been directly applied to the context of dryland rice farming in Indonesia, especially in areas with limited access such as South Papua. Furthermore, previous research tends to separate predictive [22], [23] and segmentative [24], [25], approaches, even though integrating the two, including the use of various machine learning algorithms, can provide a more comprehensive picture for decision-making.

This study offers an integrative approach which combines yield gap prediction by using a supervised learning algorithm and agricultural area segmentation through a clustering method based on input efficiency and actual yields. This approach allows for the classification of farmer groups based on their success in achieving productivity targets and the identification of areas with homogeneous input-output patterns. The newest finding of this study can be seen in the use of input efficiency ratios (seeds, pesticides, and herbicides per hectare) as the main variable in regional grouping, as well as the integration of geographic coordinate data, which opens up the potential for advanced spatial analysis.

The urgency of this research is growing as climate variability directly impacts dryland agricultural systems. Without analytical tools which are adaptive and Data-Driven intervention, the programs are vulnerable to inaccurate targeting and potentially lead to budget inefficiencies. On the other hand, a Data-Driven approach can be a solution for strengthening early warning systems, planning the distribution of superior seeds, and developing fairer and more effective agricultural insurance programs [3], [26], [27]. Local governments and agricultural policymakers can utilize the results of this study to develop evidence-based agricultural development strategies (evidence-based agricultural planning).

The primary contribution of this research is the development of a machine learning-based farmer group performance analysis framework that can be replicated in other regions with similar characteristics. This study also addresses a gap in previous research which has not simultaneously integrated predictive analysis and efficiency segmentation. Using real-world data from Merauke Regency, South Papua Province, this approach demonstrates direct application in the context of agricultural development in eastern Indonesia, which has historically lagged behind in the adoption of precision agriculture technology. Therefore, this research not only contributes to the growing scientific literature in agricultural informatics and precision agriculture but also provides practical solutions for data-driven decision-making in the agricultural sector. The implications of this research are expected to support the formulation of more adaptive, sustainable, and pro-smallholder agricultural development policies operating in areas with limited access to technology and intensive mentoring. From an applied informatics perspective, this study contributes to the development of a data-driven decision-support framework that integrates predictive analytics and clustering techniques into a unified workflow. The framework enables agricultural stakeholders to identify productivity gaps, classify farmer-group performance, and prioritize interventions using evidence-based information.

RESEARCH METHODS

This research uses a quantitative data mining-based approach by applying supervised and unsupervised machine learning techniques to analyze the performance of upland rice farming in the

ISSN 2580-6424 (printed) | ISSN 2477-2399 (online)

rainfed area of Merauke Regency, South Papua Province. The method used includes five main stages, and they are such as: (1) data collection and exploration, (2) preprocessing and feature engineering, (3) predictive modeling for the yield gap, (4) farmer group success classification, and (5) area segmentation by using clustering. This research is Data-Driven and has purpose to develop a replicable framework for Data-Driven decision-making.

Data Source and Description

The data used is the dataset of PFPL (Prospective Farmers and Prospective Locations) Upland Rice in 2025, which was obtained from technical agencies in Merauke Regency, South Papua Province, Indonesia. The dataset contains information on farmer-group locations, land area, input usage, and productivity indicators of farmer groups (sub-districts, villages, and GAPOKTAN), the identity of the group leader, land area, seed volume, dominant pest species, land type, and actual (existing) and target (target) harvest yields. The total is consisted of 30 farmer group entries spread across the rainfed area. Although the dataset size is relatively small, it represents all available PFPL upland rice farmer groups within the study area. Therefore, the analysis should be considered exploratory and intended to demonstrate the feasibility of applying machine learning techniques to local agricultural decision-support processes.

Pre-processing and Feature Engineering

The initial stage of analysis was carried out by cleaning the data from duplicates, missing values, and irrelevant entries. After that, a feature engineering process was performed to create derived variables that represent the efficiency of input use in relation to production output. The new features generated include:

1) *seed_ratio_per_ha* = seed volume (kg) / land area (Ha)
2) *pesticide_ratio_per_ha* = pesticide quantity (kg/Ltr) / land area (Ha)
3) *herbicide_ratio_per_ha* = herbicide quantity (Ltr) / land area (Ha)
4) *gap_provitas_kw_ha* = target productivity – existing productivity

The *gap_provitas_kw_ha* variable is used as the target in predictive and classification modeling.

Prediction Modeling of Yield Gap

To predict the difference between target and actual productivity (productivity gap), The Random Forest Regressor was configured with n_estimators = 100, random_state = 42. The data was separated into training and test data with a ratio of 80:20. The model was evaluated by using Root Mean Square Error (RMSE) and R^2 Score metrics to assess accuracy and the proportion of variation explained by the model. Input features used included: land area, seed volume, pesticides, herbicides, and efficiency ratio per hectare. Due to the limited dataset size, the model was evaluated using a train-test split strategy. Future studies should incorporate k-fold cross-validation to obtain a more robust estimate of model generalizability.

Classification of Successful vs. Unsuccessful Groups

Classification labels were determined based on the productivity gap criterion: if the gap was ≤ 0.5, it was categorized as successful; if it was > 0.5, it was unsuccessful. The Random Forest Classifier algorithm was used to build the classification model. Performance evaluation was conducted by examining the confusion matrix, precision, recall, and f1-score for each class. The aim of this stage was to identify the characteristics of farmer groups that were able to approach the target harvest yield.

Regional Clustering Based on Efficiency and Results

In the next step, the regional segmentation was performed by using K-Means Clustering method to group farmer groups based on input efficiency and existing productivity. The features used include: *seed_ratio_per_ha*, *pesticide_ratio_per_ha*, *herbicide_ratio_per_ha*, and *existing_provitas_kw_ha*. Data were standardized by using Standard Scaler. The optimal number of clusters was determined by using Elbow method. The clustering results were then analyzed to identify patterns of high and low efficiency areas as a basis for spatial intervention strategies.

ISSN 2580-6424 (printed) | ISSN 2477-2399 (online)

Visualization and Interpretation of Results

The prediction, classification, and clustering results were analyzed descriptively and visualized in tables and graphs to clarify differences between groups. Interpretation was performed to identify groups with high efficiency but not yet achieving targets, as well as areas with potential for best practices. These results were then used as the basis for developing Data-Driven policy recommendations.

Tools and Environment Analysis

The entire analysis process was conducted by using Python platform of 3.11 with the pandas library, scikit-learn, and matplotlib. Orange Data Mining was used for initial exploration, while cluster visualization was performed in matplotlib and seaborn. This research follows the principle of open replication (*reproducibility*) with code documentation and a pipeline which can be adapted to various regions.

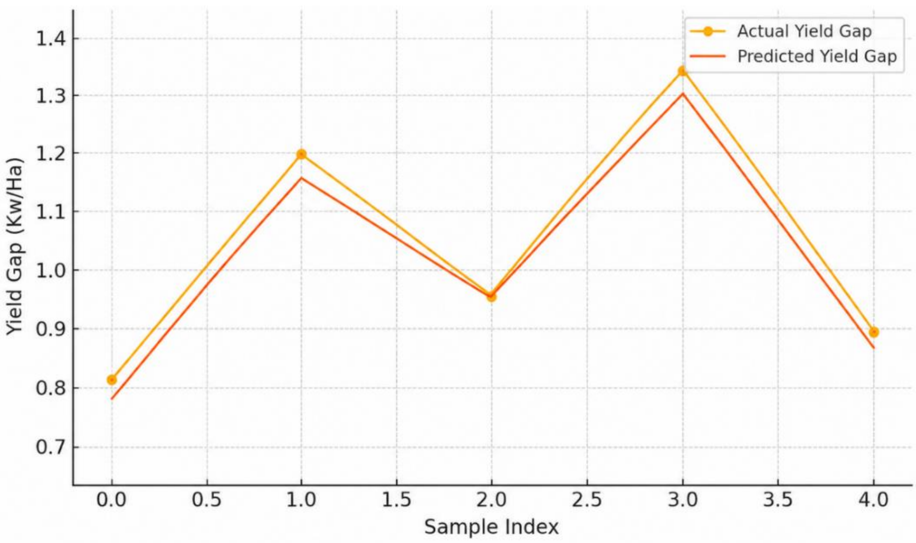
RESULT AND DISCUSSION

Result

This research produced three main interrelated outcomes: yield gap prediction, farmer group success classification, and regional efficiency segmentation by using the clustering method. Random Forest Regressor model used to predict the difference between target and actual harvest yields performed very well. The Random Forest regression model achieved a coefficient of determination (R^2) of 0.95 and an RMSE of 0.41 on the test set, While these results indicate strong predictive performance, they should be interpreted cautiously because the dataset contained only 30 observations and was evaluated using a single train-test split. Although the model demonstrated strong predictive performance, the result should be interpreted cautiously because the dataset consisted of only 30 observations, the result should be interpreted cautiously because the dataset consisted of only 30 observations. Additional validation using larger datasets and cross-validation procedures is required before claiming broader predictive capability. Figure 1. shows a close agreement between actual and predicted yield-gap values, indicating strong predictive performance with minor prediction errors.

Table 1. Metric Score

Metric	Values
R^2 Score	0.95
RMSE	0.41
MAE	0.31



ISSN 2580-6424 (printed) | ISSN 2477-2399 (online)

Figure 1. Comparison of Actual vs Predicted Yield Gaps

Farmer groups were classified based on their success in achieving productivity targets using a productivity gap_provitas_kw_ha ≤ 0.5 as an indicator of success. A Random Forest Classifier model was applied to classify groups into successful and unsuccessful classes. However, the classification results showed that all test samples fell into one class (unsuccessful), this outcome indicates severe class imbalance and limits the interpretability of classification metrics. Consequently, the classification model should be viewed as a preliminary exploratory analysis rather than a validated predictive classifier. This also indicates that the majority of farmer groups were unable to achieve their established harvest targets.

Efficiency segmentation was performed by using K-Means Clustering method with four main variables: seed_ratio_per_ha, pesticide_ratio_per_ha, herbicide_ratio_per_ha, and existing_provitas_kw_ha. To determine the optimal number of clusters, Elbow method was used, as shown in Figure 2. The elbow point is clearly visible at k = 2, which was chosen as the most appropriate number of clusters for data representation.

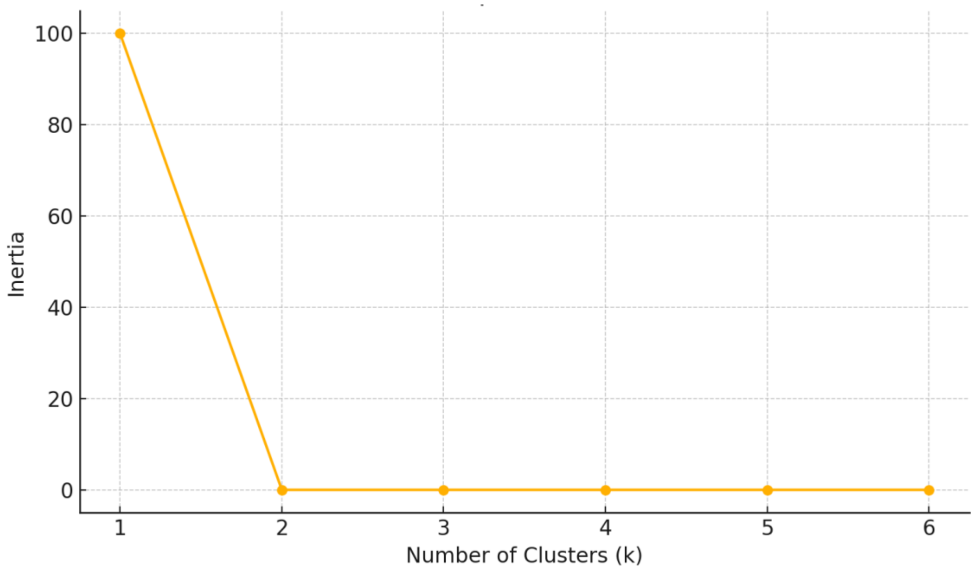


Figure 2. Elbow Method for Optimal Number of Clusters

After the clustering process, two clusters with distinct characteristics were obtained. The first cluster reflects a group of farmers with conventional practices: high inputs but relatively low yields. Conversely, the second cluster represents a group that is efficient in input use and produces much higher productivity. The average efficiency and yield variables for each cluster are shown in Table 2. below.

Table 2. Average Efficiency Ratios and Yield by Cluster

Cluster	Land Area (Ha)	Seed Ratio (kg/Ha)	Pesticide Ratio	Herbicide Ratio	Yield (Kw/Ha)
0	45.83	40.00	1.00	2.00	2.5
1	9.00	1.22	1.44	1.66	17.0

A visualization of the comparison between clusters is presented in Figure 3, highlighting the contrast between the efficient and conventional groups.

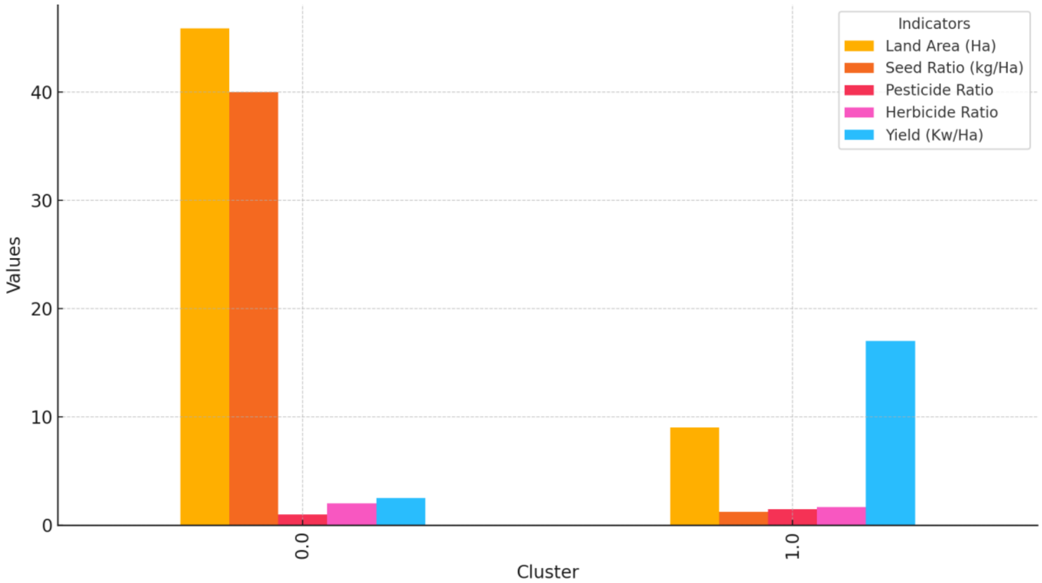


Figure 3. highlighting the contrast between the efficient and conventional groups.

Discussion

The results of this study demonstrate that a machine learning-based predictive approach can provide an accurate picture of the yield gap in upland rice farming systems in rainfed areas such as Merauke Regency, South Papua. Random Forest Regressor Model used demonstrated very high performance, with an R^2 value of 0.95 and a prediction error of nearly zero (RMSE = 0.41). This success reflects the robust structure of PFPL (Prospective Farmers and Prospective Locations) data used to model the relationship between input variables (land area, seeds, pesticides, herbicides) and productivity output. These findings are consistent with previous studies demonstrating that machine learning models can address the complexity of agricultural data with superior predictive results [28], [29].

However, the classification of farmer group success using Random Forest Classifier algorithm demonstrated suboptimal results due to the dominance of the "unsuccessful" class in the data. This imbalance in label distribution indicates that the majority of groups have not been able to achieve their set productivity targets. This clearly reflects the challenges faced in marginalized areas like South Papua, where environmental conditions, access to technology, and farmers' technical skills remain major obstacles. On the other hand, this situation opens up opportunities for further research with additional data from successful farmer groups to refine the classification model and support a more representative machine learning process.

The clustering process resulted in two groups of farming areas with starkly contrasting characteristics. The first cluster represents common agricultural practices such as the use which was high but inefficient input with relatively low yields. In contrast, the second cluster demonstrates better input efficiency and produces significantly higher productivity. Interestingly, based on the clustering results, groups in the efficient cluster actually have a very low seed ratio, yet manage to achieve productivity of up to 17 kW/ha. This finding suggests that efficiency, not input quantity, is the key to success in dryland farming systems [30], [31].

The policy implications of this research are important for local governments and policymakers in the agricultural sector to consider. First, The findings may provide preliminary guidance for local governments in designing cluster-based intervention strategies, where groups in the efficient cluster serve as models of best practices and receive support to become training centers for other farmers. Meanwhile, groups in low-efficiency clusters need to be provided with intensive mentoring, including training in cultivation techniques, reformulating input distribution, and evaluating the effectiveness of more targeted subsidies. Data-Driven, transdisciplinary policies like this would be more adaptive than the uniform national approach currently implemented without taking local variations into account [32], [33].

Second, the results of this analysis can also form the basis for planning performance-based incentive programs, where farmer groups with high efficiency and consistent yields are provided with

ISSN 2580-6424 (printed) | ISSN 2477-2399 (online)

additional support such as market access, agricultural insurance programs, and farm credit. Furthermore, with the geographic features (coordinate points) in PFPL data, a spatial-based information system can be built as the basis for a decision support system (DSS) for the agricultural sector. This approach aligns with the national vision of agricultural digitalization and simultaneously supports the Sustainable Development Goals (SDGs) for food security, poverty alleviation, and increased agricultural productivity in disadvantaged regions [34], [35].

CONCLUSION

This study demonstrates that PFPL (Prospective Farmer Prospective Location) Data-Driven machine learning approach can provide a comprehensive understanding of the yield gap and input efficiency in upland rice farming systems in the rainfed area of Merauke Regency, South Papua Province. The regression model produced highly accurate predictions of the gap between target and realized productivity, while the clustering process successfully identified two main groups, and they are such as: conventional farmer groups with high inputs but low yields, and efficient groups with low inputs but high productivity. These results confirm that the efficiency of agricultural input use, not simply the quantity, is the primary determinant of achieving optimal yields in dryland areas. These findings may provide preliminary insights for efficiency-based agricultural interventions and data-driven decision-support systems, particularly for designing efficiency-zoning-based interventions and developing more contextual and Data-Driven decision-making systems.

For future research, the integration of supporting variables such as climate data, soil conditions, and cultivation practices is recommended to increase the model's reach across various agro-ecological contexts. This study is limited by the relatively small sample size (30 farmer groups), class imbalance in the classification stage, and the use of a single train-test split for model evaluation. Therefore, the findings should be interpreted as exploratory and require validation using larger and more diverse datasets. Expanding the sample size from various regions and farmer groups that have achieved productivity targets is also crucial to address class imbalance in the classification model. Furthermore, the integration of spatial and remote sensing data has the potential to strengthen spatial analysis and area-based policy mapping. The development of an interactive dashboard system based on these findings will be highly beneficial for local governments and agricultural extension workers in developing more adaptive and sustainable food security policies and performance-based interventions.

ACKNOWLEDGMENT

The author would like to thank the Department of Agriculture, Food Crops, and Horticulture of Merauke Regency, South Papua, for providing valid data, enabling the author to complete this article successfully.

REFERENCES

[1] U. Faoziyah, M. F. Rosyaridho, and R. Panggabean, "Unearthing Agricultural Land Use Dynamics in Indonesia: Between Food Security and Policy Interventions," *Land (Basel)*, vol. 13, no. 12, p. 2030, Nov. 2024, doi: 10.3390/land13122030.

[2] S. Wang *et al.*, "Rice yield stability and its determinants across different rice-cropping systems in China," *Agric For Meteorol*, vol. 364, p. 110452, Apr. 2025, doi: 10.1016/j.agrformet.2025.110452.

[3] A. El Mane, Y. Chihab, K. Tatane, and R. Korchiyne, "Agriculture Supply Chain Management Based on Blockchain Architecture and Smart Contracts," *Applied Computational Intelligence and Soft Computing*, vol. 2022, pp. 1–23, Oct. 2022, doi: 10.1155/2022/8011525.

[4] S. Raodah, M. Nur, and M. S. Lestari, "Jurus Jitu Tingkatkan Produksi Padi Sawah Tadah Hujan," *KEMENTERIAN PERTANIAN RI*, 2023.

[5] R. Raza, R. Director, I. Cossio, and C. Country, "Technical questions: International Fund for Agricultural Development-Executive Board Republic of Indonesia Country strategic opportunities programme," Dec. 2023. [Online]. Available: www.ifad.org

[6] R. Tirtalistyani, M. Murtiningrum, and R. S. Kanwar, "Indonesia Rice Irrigation System: Time for Innovation," *Sustainability*, vol. 14, no. 19, p. 12477, Sep. 2022, doi: 10.3390/su141912477.

[7] I. P. N. Damanik, M. E. Tahitu, and E. Kembauw, "Analysis of Farmer's Group Functions in the Adoption Process of Superior Seed Innovation in Waimital Village, Kairatu District," *International Journal of Multidisciplinary Sciences and Arts*, vol. 1, no. 2, pp. 140–146, Jan. 2023, doi: 10.47709/ijmdsa.v1i2.2017.

[8] S. M. Shawon, F. B. Ema, A. K. Mahi, F. L. Niha, and H. T. Zubair, "Crop yield prediction using machine learning: An extensive and systematic literature review," *Smart Agricultural Technology*, vol. 10, p. 100718, Mar. 2025, doi: 10.1016/j.atech.2024.100718.

[9] R. Gupta, T. Shrikant Padmawar, D. Kumar, D. Ray, and P. Kadam, "Significance of Machine Learning in Crop Yield Prediction," in *2024 2nd World Conference on Communication & Computing (WCONF)*, IEEE, Jul. 2024, pp. 1–7. doi: 10.1109/WCONF61366.2024.10692141.

[10] Md. A. Jabed and M. A. Azmi Murad, "Crop yield prediction in agriculture: A comprehensive review of machine learning and deep learning approaches, with insights for future research and sustainability," *Heliyon*, vol. 10, no. 24, p. e40836, Dec. 2024, doi: 10.1016/j.heliyon.2024.e40836.

[11] K. Ajaykumar and S. Madhavi, "Review on Crop Yield Prediction with Deep Learning and Machine Learning Algorithms," in *2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA)*, IEEE, Sep. 2022, pp. 903–909. doi: 10.1109/ICIRCA54612.2022.9985016.

[12] P. Schatte and M. Meyer, "Assessing holistic agroecological resilience of agroecosystems from a landscape perspective: a systematic review," *Ecology and Society*, vol. 30, no. 2, p. art24, 2025, doi: 10.5751/ES-16137-300224.

[13] J. Botero-Valencia *et al.*, "Machine Learning in Sustainable Agriculture: Systematic Review and Research Perspectives," *Agriculture*, vol. 15, no. 4, p. 377, Feb. 2025, doi: 10.3390/agriculture15040377.

[14] L. Fan, S. Fang, J. Fan, Y. Wang, L. Zhan, and Y. He, "Rice Yield Estimation Using Machine Learning and Feature Selection in Hilly and Mountainous Chongqing, China," *Agriculture*, vol. 14, no. 9, p. 1615, Sep. 2024, doi: 10.3390/agriculture14091615.

[15] E. A. Jiya, U. Illiyasu, and M. Akinyemi, "Rice Yield Forecasting: A Comparative Analysis of Multiple Machine Learning Algorithms," *Journal of Information Systems and Informatics*, vol. 5, no. 2, pp. 785–799, Jun. 2023, doi: 10.51519/journalisi.v5i2.506.

[16] T. Islam, T. Mazumder, Md. N. Shahriair Roni, and Md. S. Nur, "A comparative study of machine learning models for predicting Aman rice yields in Bangladesh," *Heliyon*, vol. 10, no. 23, p. e40764, Dec. 2024, doi: 10.1016/j.heliyon.2024.e40764.

[17] P. Sharma, P. Dadheech, N. Aneja, and S. Aneja, "Predicting Agriculture Yields Based on Machine Learning Using Regression and Deep Learning," *IEEE Access*, vol. 11, pp. 111255–111264, 2023, doi: 10.1109/ACCESS.2023.3321861.

[18] S. Iniyan, V. Akhil Varma, and C. Teja Naidu, "Crop yield prediction using machine learning techniques," *Advances in Engineering Software*, vol. 175, p. 103326, Jan. 2023, doi: 10.1016/j.advengsoft.2022.103326.

[19] S. Sow *et al.*, "Artificial Intelligence for Maximizing Agricultural Input Use Efficiency: Exploring Nutrient, Water and Weed Management Strategies," *Phyton (B Aires)*, vol. 93, no. 7, pp. 1569–1598, 2024, doi: 10.32604/phyton.2024.052241.

[20] S. O. Araújo, R. S. Peres, J. C. Ramalho, F. Lidon, and J. Barata, "Machine Learning Applications in Agriculture: Current Trends, Challenges, and Future Perspectives," *Agronomy*, vol. 13, no. 12, p. 2976, Dec. 2023, doi: 10.3390/agronomy13122976.

[21] B. Dey, J. Ferdous, and R. Ahmed, "Machine learning based recommendation of agricultural and horticultural crop farming in India under the regime of NPK, soil pH and three climatic variables," *Heliyon*, vol. 10, no. 3, p. e25112, Feb. 2024, doi: 10.1016/j.heliyon.2024.e25112.

[22] K. Singh, M. Yadav, D. Barak, S. Bansal, and F. Moreira, "Machine-Learning-Based Frameworks for Reliable and Sustainable Crop Forecasting," *Sustainability*, vol. 17, no. 10, p. 4711, May 2025, doi: 10.3390/su17104711.

[23] G. Mohyuddin, M. A. Khan, A. Haseeb, S. Mahpara, M. Waseem, and A. M. Saleh, "Evaluation of Machine Learning Approaches for Precision Farming in Smart Agriculture System: A Comprehensive Review," *IEEE Access*, vol. 12, pp. 60155–60184, 2024, doi: 10.1109/ACCESS.2024.3390581.

[24] K. P. Swain *et al.*, "Empowering Crop Selection with Ensemble Learning and K-means Clustering: A Modern Agricultural Perspective," *Open Agric J*, vol. 18, no. 1, Feb. 2024, doi: 10.2174/0118743315291367240207093403.

[25] A. J. Rivera, M. D. Pérez-Godoy, D. Elizondo, L. Deka, and M. J. del Jesus, "Analysis of clustering methods for crop type mapping using satellite imagery," *Neurocomputing*, vol. 492, pp. 91–106, Jul. 2022, doi: 10.1016/j.neucom.2022.04.002.

[26] C. P. Osorio, F. Leucci, and D. Porrini, "Analyzing the Relationship between Agricultural AI Adoption and Government-Subsidized Insurance," *Agriculture*, vol. 14, no. 10, p. 1804, Oct. 2024, doi: 10.3390/agriculture14101804.

[27] L. Hu, C. Zhang, M. Zhang, Y. Shi, J. Lu, and Z. Fang, "Enhancing FAIR Data Services in Agricultural Disaster: A Review," *Remote Sens (Basel)*, vol. 15, no. 8, p. 2024, Apr. 2023, doi: 10.3390/rs15082024.

[28] M. Waqas, A. Naseem, U. W. Humphries, P. T. Hlaing, P. Dechpichai, and A. Wangwongchai, “Applications of machine learning and deep learning in agriculture: A comprehensive review,” *Green Technologies and Sustainability*, vol. 3, no. 3, p. 100199, Jul. 2025, doi: 10.1016/j.grets.2025.100199.

[29] D. Tamayo-Vera, X. Wang, and M. Mesbah, “A Review of Machine Learning Techniques in Agroclimatic Studies,” *Agriculture*, vol. 14, no. 3, p. 481, Mar. 2024, doi: 10.3390/agriculture14030481.

[30] Y. Xing and X. Wang, “Precision Agriculture and Water Conservation Strategies for Sustainable Crop Production in Arid Regions,” *Plants*, vol. 13, no. 22, p. 3184, Nov. 2024, doi: 10.3390/plants13223184.

[31] L. Xiong and W. Wu, “Can additional agricultural resource inputs improve maize yield, resource use efficiencies and emergy based system efficiency under ridge-furrow with plastic film mulching?,” *J Clean Prod*, vol. 379, p. 134711, Dec. 2022, doi: 10.1016/j.jclepro.2022.134711.

[32] L. Sweet *et al.*, “Transdisciplinary coordination is essential for advancing agricultural modeling with machine learning,” *One Earth*, vol. 8, no. 4, p. 101233, Apr. 2025, doi: 10.1016/j.oneear.2025.101233.

[33] E. Grigorieva, A. Livenets, and E. Stelmakh, “Adaptation of Agriculture to Climate Change: A Scoping Review,” *Climate*, vol. 11, no. 10, p. 202, Oct. 2023, doi: 10.3390/cli11100202.

[34] A. M. Hiywotu, “Advancing sustainable agriculture for goal 2: zero hunger - a comprehensive overview of practices, policies, and technologies,” *Agroecology and Sustainable Food Systems*, pp. 1–29, Jan. 2025, doi: 10.1080/21683565.2025.2451344.

[35] A. J. Atapattu, C. S. Ranasinghe, T. D. Nuwarapaksha, S. S. Udumann, and N. S. Dissanayaka, “Sustainable Agriculture and Sustainable Development Goals (SDGs),” 2024, pp. 1–27. doi: 10.4018/979-8-3693-4864-2.ch001.