Automatically Retrained Machine Learning System for Rice Yield Prediction Using Open-Meteo and BPS Data

Ulfatun Nadifa*, Ikhsan Hidayat, Wildan

Department of Electrical and Computer Engineering, Faculty of Engineering, Universitas Negeri Gorontalo
No. 6, Jenderal Sudirman Street, Gorontalo City, Gorontalo Province, Indonesia
**Corresponding author. Email: ulfatun@ung.ac.id*

Abstract— This study proposes a machine learning–based rice yield prediction system with a self-updating mechanism, using Gorontalo Province, Indonesia, as a case study. The system integrates daily climate data from Open-Meteo with agricultural statistics from the Central Bureau of Statistics (BPS) to support data-driven decision-making in agriculture. A key challenge addressed in this study is the limited availability of yield data, which are provided only at an annual scale for the period 2018–2024, without seasonal labels. To overcome this limitation, a temporal disaggregation approach is adopted to construct initial seasonal yield labels (M1, M2, M3). These constructed labels serve as approximations, enabling the development of a seasonal prediction model under data-constrained conditions. Several machine learning algorithms, namely Gradient Boosting, Random Forest, XGBoost, Ridge Regression, and Linear Regression, are evaluated using Leave-One-Out Cross-Validation (LOO-CV). Model performance is assessed using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2). The results indicate that ensemble-based models outperform linear baselines, with Gradient Boosting providing the best balance between prediction accuracy and model stability. The main contribution of this study is the design of an adaptive prediction system with a self-updating mechanism that supports periodic retraining and dynamic model evaluation. At its current stage, this mechanism is positioned as an initial framework rather than a fully validated continuous learning system. The proposed system is implemented as a web-based platform that supports yield prediction and planting season recommendations, providing a scalable foundation for intelligent agricultural systems in data-limited environments.

Keywords— Automatic retraining; BPS; Climate data; Gradient boosting; Machine learning; Web-based decision support system

I. INTRODUCTION

Rice is one of the primary staple crops that plays a strategic role in ensuring national food security. However, the agricultural sector is increasingly challenged by global climate change, including the rising frequency of extreme weather events and unpredictable shifts in planting seasons [1]. These conditions introduce significant uncertainty into crop yields and pose a serious threat to agricultural productivity. Therefore, accurate, adaptive, and data-driven crop yield prediction systems are essential to support decision-making in planting schedules, land management, and sustainable food security policies.

In recent years, Machine Learning (ML) and Artificial Intelligence (AI) approaches have been widely adopted for crop yield prediction owing to their ability to capture complex and nonlinear relationships between climatic variables and crop production [2]. Previous studies have shown that climate factors such as temperature, rainfall, humidity, and solar radiation significantly influence rice yield variability [1], [3]. Furthermore, the integration of climate data with feature selection techniques and remote sensing information has been shown to markedly improve prediction accuracy compared with conventional statistical methods [4]. Ensemble and tree-based models such as Random Forest [5], XGBoost, and LightGBM have demonstrated strong performance in modeling nonlinear relationships, while advances in deep learning indicate a shift toward more intelligent and data-driven agricultural systems [6], [7].

Several studies have further explored the relationship between climatic variables and rice yield. Tasneem et al. [1] highlighted the impact of extreme temperature variability on yield fluctuations, while Roy et al. [3] identified rainfall and minimum temperature as key influencing factors. Fan et al. [4] demonstrated that combining climate data with feature selection and topographic information improves prediction accuracy, whereas Sah et al. [8] showed that integrating biophysical and remote sensing data leads to more stable predictions. In addition, Meroni et al. [9] confirmed that reliable crop yield prediction can still be achieved under limited data conditions when appropriate modeling strategies are applied.

Nevertheless, most existing studies rely on large, complete, and static historical datasets that are assumed to be fully available at the initial stage of model development. Moreover, these approaches generally adopt a static training paradigm, in which models are trained only once and lack mechanisms for continuous updating when new data become available. As a result, such models exhibit limited adaptability to evolving climate conditions, seasonal variability, and dynamic yield patterns. In addition, although climate data are widely used, the utilization of open-access climate services such as Open-Meteo

Received 19 January 2026, Revised 11 April 2026, Accepted 15 April 2026.
DOI: https://doi.org/10.15294/jte.v18i1.41368

as a primary data source remains limited, particularly in developing regions where access to local meteorological observations is constrained [10].

Another challenge lies in the mismatch between data sources. Agricultural data provided by the Central Bureau of Statistics (BPS) are typically available only as annual aggregates, whereas climate data are available at a much higher temporal resolution (e.g., daily). This discrepancy makes it difficult to develop season-based prediction models. To address this issue, this study adopts a temporal disaggregation approach, in which annual yield data are distributed across multiple planting seasons based on climatic patterns, enabling the construction of initial seasonal labels that can be refined over time.

It should be noted that the seasonal yield labels constructed in this study are derived through a temporal disaggregation approach due to the absence of real seasonal observations. Therefore, the generated labels represent an initial approximation rather than exact seasonal measurements. This approach is intended to enable model development under data-constrained conditions and can be progressively refined as more detailed data become available.

Based on these gaps, this study proposes a self-updating rice yield prediction system designed to operate under limited initial data conditions while continuously improving through periodic retraining as new data become available. The system integrates daily climate data from Open-Meteo as an open-access data source and supports adaptive prediction under changing climate conditions. Through this approach, the proposed system not only enhances prediction capability but also establishes a foundation for continuously evolving, data-driven agricultural decision support systems. At its current stage, the proposed self-updating mechanism is positioned as an initial adaptive framework.

II. Method

This study adopts an experimental machine learning–based approach integrated with the development of a self-updating artificial intelligence system for seasonal rice yield prediction. The overall research workflow (Figure 1) comprises data collection, data preprocessing, feature construction, model development, system implementation, and continuous model updating.

A. Data Sources and Preprocessing

The data used in this study consist of daily climate variables obtained from the open-access climate data service Open-Meteo and rice yield data sourced from the BPS, which serve as the initial training labels. Open-Meteo was selected due to its capability to provide high-resolution and near real-time climate information, making it well suited for continuously updated prediction systems and web-based forecasting platforms [11], [12], [13].

Since the rice yield data obtained from BPS are only available in annual aggregates, a temporal disaggregation approach was applied to construct seasonal yield labels. The annual yield data were distributed into three planting seasons based on climatic patterns and cropping cycles, enabling consistency between high-resolution climate inputs and seasonal prediction targets.

The climate and yield datasets were subsequently integrated by aligning temporal information at the seasonal level (Table 1). The preprocessing stage involved data cleaning, handling missing values, normalization, and aggregation of daily climate variables into seasonal indicators. This process is intended to capture the relationship between climatic conditions and seasonal rice yield [14].

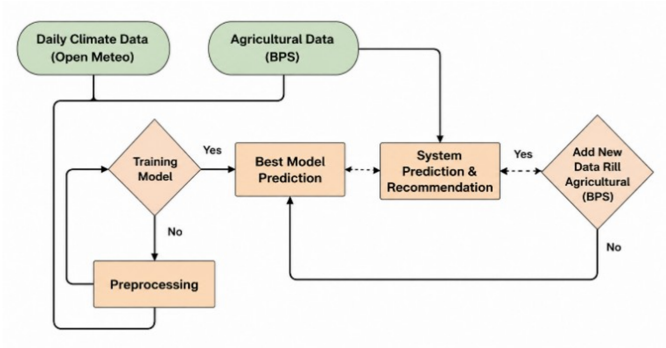

Figure 1. Flowchart of the System for Seasonal Rice Yield Prediction

TABLE I. INPUT FEATURES USED IN THE PREDICTION MODEL

No	Feature Name	Description
1	precipitation_sum	Total precipitation (mm) accumulated over a specific time period.
2	rain_sum	Total liquid rainfall (mm), excluding snow.
3	showers_sum	Total heavy shower-type rainfall (mm).
4	snowfall_sum	Total snowfall converted into water equivalent (mm); generally zero in tropical regions.
5	shortwave_radiation_sum	Total shortwave solar radiation received at the Earth's surface (MJ/m^2).
6	sunshine_duration	Total duration of sunshine exposure (seconds or hours per day).
7	et0_fao_evapotranspiration	Reference evapotranspiration (ET_0) based on the FAO method, representing plant water demand.
8	temperature_2m_mean	Mean air temperature at 2 meters above ground level (°C).
9	temperature_2m_max	Maximum daily air temperature at 2 meters above ground level (°C).
10	temperature_2m_min	Minimum daily air temperature at 2 meters above ground level (°C).
11	relative_humidity_2m_mean	Mean relative humidity at 2 meters above ground level (%).
12	relative_humidity_2m_max	Maximum daily relative humidity at 2 meters above ground level (%).
13	relative_humidity_2m_min	Minimum daily relative humidity at 2 meters above ground level (%).
14	wind_speed_10m_mean	Mean wind speed at 10 meters above ground level (m/s).
15	wind_speed_10m_max	Maximum wind speed at 10 meters above ground level (m/s).
16	wind_speed_10m_min	Minimum wind speed at 10 meters above ground level (m/s).
17	wind_direction_10m_dominant	Dominant wind direction at 10 meters above ground level (degrees).
18	cloud_cover_mean	Average percentage of cloud cover in the sky (%).
19	soil_temperature_0_to_7cm_mean	Mean soil temperature at a depth of 0–7 cm (°C).
20	soil_moisture_0_to_7cm_mean	Mean soil moisture content at a depth of 0–7 cm (m^3/m^3).
21	Land_area	Cultivated land area used for agricultural production (hectares).

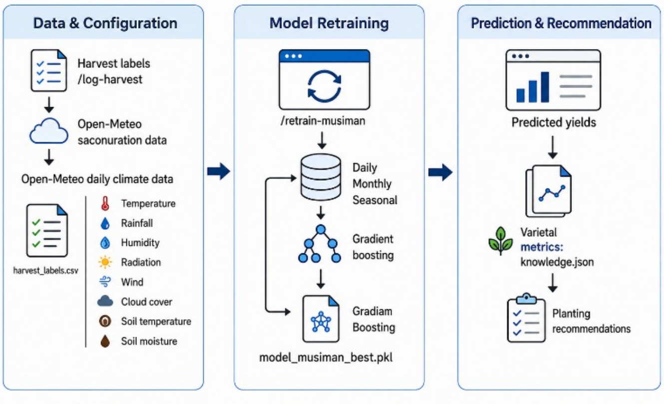

Figure 2. Self-updating Seasonal Rice Yield Prediction Pipeline

B. Feature Construction and Variable Selection

The features used in this study consist of seasonal climate variables, including temperature, rainfall, humidity, solar radiation, wind speed, and other parameters related to atmospheric and soil conditions. In addition, non-climatic factors such as cultivated land area are included as supporting variables in the model. To ensure model validity and prevent target leakage, variables that directly represent crop yield or productivity (e.g., productivity or yield per hectare) are excluded from the input feature set and are used solely as prediction targets.

To reduce feature redundancy and enhance model stability, feature selection plays an important role in machine learning-based prediction systems. One commonly used approach is Least Absolute Shrinkage and Selection Operator (LASSO)-ased regularization, which is effective in identifying the most influential variables by shrinking less important coefficients toward zero, thereby improving model interpretability and generalization performance [15].

However, LASSO has certain limitations. It tends to select only one variable among highly correlated features, potentially ignoring other relevant predictors. Moreover, its linearity assumption may limit its ability to capture complex nonlinear relationships between climate variables and crop yield. Therefore, in this study, nonlinear ensemble-based models are prioritized to better capture complex interactions among variables.

C. Seasonal Label Construction

To address the absence of seasonal yield data, a temporal disaggregation approach was employed to transform annual rice yield into seasonal yield labels. The annual yield (Ya) was distributed into three planting seasons (M1, M2, and M3) using proportional weighting, as expressed in Equation (1):

$$Y_s = w_s \times Y_a \quad (1)$$

where Ys represents the estimated yield for a given season, Ya denotes the total annual yield, and ws is the seasonal weight. The seasonal weights were constrained such that their total sum equals one, as defined in Equation (2):

$$\sum_{s=1}^{3} w_s = 1 \quad (2)$$

The seasonal weights (w_s) were determined based on rainfall distribution patterns and regional cropping calendars, reflecting the relative contribution of each planting season to total annual production. Rainfall was used as the primary proxy variable due to its strong influence on rice growth and productivity, as widely reported in agricultural and climate yield studies [1], [3].

This temporal disaggregation approach is commonly adopted in situations where high-resolution labels are unavailable, enabling alignment between fine-grained input features (e.g., daily or seasonal climate variables) and coarse-grained target variables (annual yield) [9]. Although the resulting seasonal labels are approximations, they provide a practical and systematic basis for model training under data-limited conditions. Accordingly, the constructed seasonal labels should be interpreted as proxy targets rather than exact observations. This assumption introduces inherent uncertainty into the constructed labels, which will be addressed in future work through the integration of real seasonal observations.

D. Model Training and Evaluation

Several machine learning algorithms were employed in this study, including Gradient Boosting, Random Forest, XGBoost, Ridge Regression, and Linear Regression. Given the limited size of the initial dataset, model evaluation was conducted using LOO-CV, which is suitable for small-sample scenarios as it maximizes data utilization by iteratively using one observation as the test set while the remaining observations are used for training [16].

Prior to model training, all input features were normalized using a scaling technique to improve training stability, particularly for linear-based algorithms. The training process was implemented using a pipeline framework that integrates feature scaling and model learning in a structured manner.

Model performance was evaluated using three primary metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2) [17]. These metrics were used to quantify prediction errors and assess the model's ability to explain the variability in the data.

All evaluation metrics were computed on the original scale (tons per hectare) after applying inverse transformation, ensuring that the results remain interpretable within the context of agricultural production.

The selection of algorithms in this study is based on their capability to capture complex relationships between climatic variables and crop yield. Ensemble-based models such as Random Forest, Gradient Boosting, and XGBoost were chosen due to their effectiveness in modeling nonlinear patterns, while linear models such as Linear Regression and Ridge Regression were included as baseline models for comparison.

E. System Implementation and Self-Updating Mechanism

The selected model was integrated into a FastAPI-based Application Programming Interface (API) and deployed within a web-based system developed using Laravel and React. The system enables user input of land-related data, automatic retrieval of current-year climate data from the Open-Meteo service, and real-time visualization of predicted rice yield along with planting season recommendations.

The system adopts a service-oriented architecture that facilitates seamless interaction between data processing modules, prediction models, and the user interface. Climate data are dynamically obtained via the Open-Meteo API and subsequently transformed into seasonal features before being used as input for the prediction model.

A key feature of the proposed system is its automatic model updating (self-updating) mechanism. After each harvest period, newly observed yield data collected from the field are stored as additional training labels and incorporated into periodic model retraining. This enables the model to adapt as new data become available, allowing gradual adjustment to evolving climate conditions and agricultural dynamics.

In addition, the system performs re-evaluation of all candidate models during each retraining cycle. This process allows for potential shifts in the best-performing model (model shifting) as the dataset grows. Rather than relying on a fixed model, the system dynamically selects the most suitable model based on updated data.

This approach supports continuous monitoring of model performance over time and provides a basis for future analysis of how changes in data distribution and climate variability may influence predictive accuracy. Consequently, the system is designed not only as a prediction tool but also as an initial framework for a continuous learning platform that supports adaptive, data-driven decision-making in agriculture [19]. At its current stage, the self-updating mechanism is positioned as an initial adaptive framework and has not yet been validated through longitudinal evaluation.

III. Results and Discussion

This section presents the experimental results and analyzes the performance of the proposed rice yield prediction system. The discussion focuses on a comparative assessment of prediction accuracy and model stability, as well as the role of the self-updating mechanism in improving model adaptability over time.

Furthermore, the contribution of real-time climate data obtained from the Open-Meteo service is examined in the context of enabling dynamic and seasonally adaptive yield predictions. The results are also analyzed to highlight their practical relevance for agricultural decision-making, particularly in supporting planting strategies and resource management.

Finally, the limitations of the proposed approach are discussed, along with potential directions for future research aimed at improving data quality, model robustness, and system scalability.

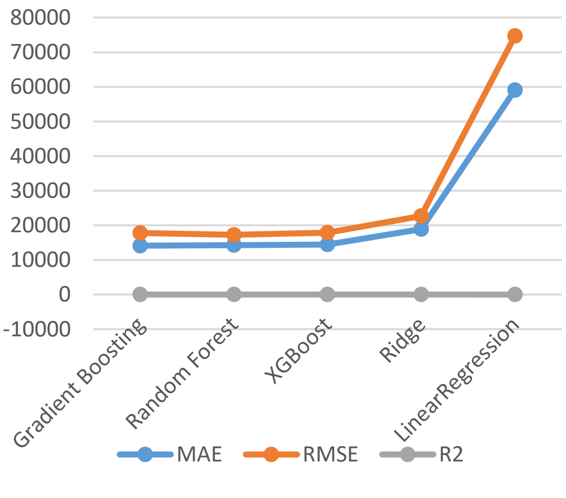

Figure 3. Comparison of Model Performance Results under LOO-CV

A. Model Performance Evaluation and Selection Model

TABLE II. Model Performance Evaluation Results (LOO-CV)

Model	MAE	RMSE	R^2
Gradient Boosting	14112	17785	0.1977
Random Forest	14249	17260	0.2443
XGBoost	14425	17853	0.1915
Ridge	18872	22644	-0.3006
Linear Regression	59045	74655	-13.1367

Model performance was evaluated using LOO-CV to address the limited size of the available dataset (Figure 3). Five machine learning algorithms were assessed, including Gradient Boosting, Random Forest, XGBoost, Ridge Regression, and Linear Regression. Model performance was measured using MAE, RMSE, and R^2 as evaluation metrics (Table II).

In this study, Linear Regression and Ridge Regression were used as baseline models to provide a reference for evaluating the performance of ensemble-based approaches. The results indicate that ensemble models, particularly Random Forest and Gradient Boosting, demonstrated more stable performance, as reflected by positive $R2$ values, indicating their better ability to capture nonlinear relationships between climate variables and rice yield [5], [21].

Although Random Forest achieved a relatively higher R^2 in some cases, it tends to exhibit greater prediction variance due to its bagging-based structure, making it more sensitive to data fluctuations in small datasets [18]. In contrast, Gradient Boosting employs a sequential learning approach that incrementally minimizes residual errors, resulting in more stable and consistent predictions [21]. The relatively lower performance of XGBoost observed in this study may be associated with its higher model complexity, which may not confer additional benefits under limited data conditions.

The baseline models showed weaker performance, indicated by negative $R2$ values, suggesting that their predictive capability is inferior to a simple mean-based baseline [17]. Although Linear Regression may yield relatively low MAE values in certain cases, higher RMSE values indicate larger error variance and reduced prediction stability. These findings highlight the importance of using multiple evaluation metrics rather than relying on a single measure alone.

To support objective model selection, the developed system, particularly within the FastAPI-based implementation, incorporates a multi-metric evaluation framework. This approach enables comprehensive assessment of model performance by simultaneously considering accuracy, stability, and generalization capability.

Based on these considerations, Gradient Boosting was selected as the operational model, as it provides a balanced trade-off between predictive accuracy and model stability. This finding is consistent with previous studies that highlight the effectiveness of boosting-based methods in modeling nonlinear relationships in agricultural data [5], [21].

B. System Performance under the Self-Updating Scheme

The primary advantage of the proposed system lies in the implementation of an automatic model updating (self-updating) mechanism (Figure 2), in which newly observed harvest data are incorporated as additional labels for periodic model retraining (Figure 4). This approach is intended to enable the model to adapt to evolving climatic conditions and changes in land productivity over time. This mechanism is consistent with adaptive learning concepts in seasonal forecasting and climate-smart agriculture systems (Figure 5) [6], [19], [22].

In addition, the system re-evaluates all candidate models during each retraining cycle, allowing for potential shifts in the best-performing model (model shifting) as part of the adaptive system design, as the dataset grows over time. This approach helps maintain the relevance of the selected operational model under changing data conditions.

However, it should be clearly noted that the self-updating mechanism in this study has not yet been evaluated through a time-based or incremental experimental framework. Specifically, the study does not include a comparative analysis between static and continuously updated models, nor does it employ rolling-origin evaluation or concept drift monitoring. Therefore, the current implementation should be regarded as an initial system-level framework, while comprehensive experimental validation of its long-term performance remains an important direction for future work.

The use of Open-Meteo as an open-access climate data source enables the system to generate dynamic predictions supported by near real-time daily data. However, its relatively coarse spatial resolution may not fully capture local microclimatic variations, highlighting the importance of integrating local observations, such as those provided by Meteorological, Climatological, and Geophysical Agency of Indonesia (BMKG), to improve prediction accuracy [23].

The prediction outputs are utilized not only for estimating rice production but also for generating planting season recommendations, enabling the system to function as a Decision Support System (DSS) for farmers and policymakers (Figure 6). This can support planting time decisions, help reduce the risk of crop failure due to extreme weather events, and facilitate strategic planning for food distribution and reserve management [10].

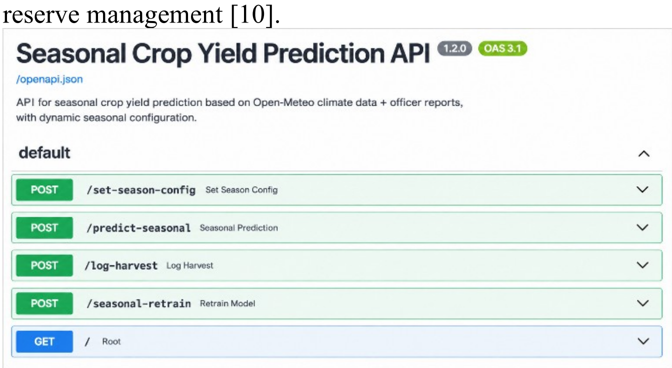

Figure 4. Seasonal Crop Yield Prediction API

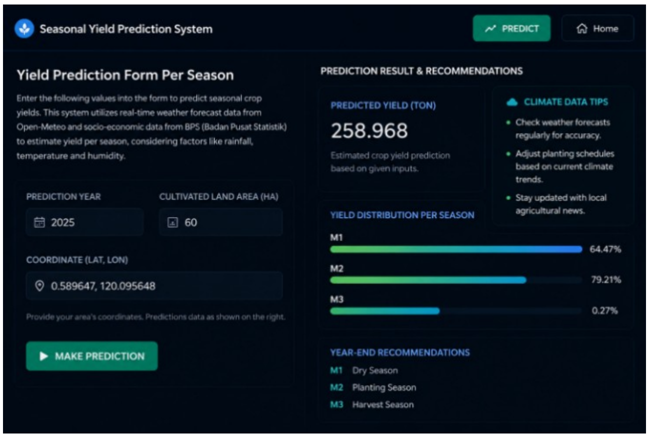

Figure 5. Seasonal Yield Prediction System

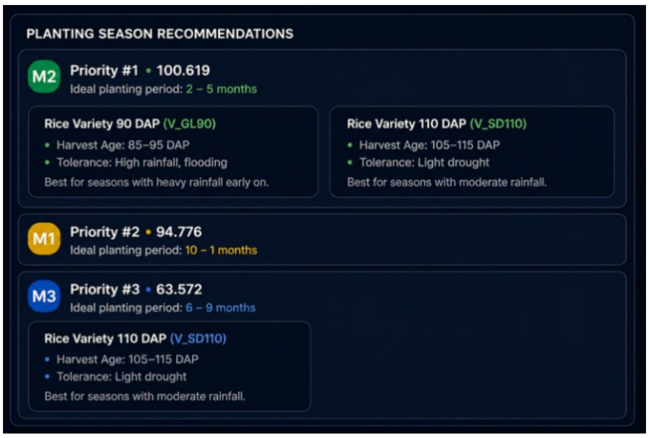

Figure 6. Planting Season Recommendation

Despite these advantages, several limitations remain, including limited historical yield data, reliance on global climate data sources, and the restriction to seasonal-scale predictions. These limitations highlight the importance of incorporating higher-resolution yield data and conducting longitudinal evaluation in future work. Future research should focus on expanding data availability, integrating local climate observations, and exploring more adaptive modelling approaches to enhance system performance.

IV. Conclusion

This study proposed a self-updating artificial intelligence–based system for seasonal rice yield prediction in Gorontalo Province, Indonesia, by integrating machine learning, climate data processing, and a service-oriented architecture. The experimental results showed that ensemble-based models, particularly Gradient Boosting, achieved more reliable and stable performance than baseline linear methods under limited data conditions. The system also demonstrated the potential of integrating open-access climate data from Open-Meteo to support adaptive and data-driven agricultural decision-making. In addition, the proposed self-updating mechanism enables periodic model retraining using newly available harvest data, although its effectiveness has not yet been validated through time-based experimental evaluation. Future research should focus on expanding dataset availability, incorporating local climate observations and IoT-based sensing systems, and evaluating more advanced learning approaches to improve model robustness, scalability, and real-world applicability.

Acknowledgment

The authors would like to express their sincere gratitude to the Central Bureau of Statistics (BPS) for providing agricultural production data and to Open-Meteo for making open-access climate data available for this research. Special appreciation is also extended to all field officers and stakeholders who contributed to data collection and system validation. This research was conducted as part of an academic and scientific effort to support the development of intelligent and adaptive agricultural prediction systems in Indonesia, particularly in Gorontalo Province.

References

[1] K. T. Tasneem, M. Umair, J. Rashid, K. M. Othman, T. Zafar, and M. Faheem, “Predicting rice yield and impact of climate change on rice production using machine learning models,” Theor. Appl. Climatol., 2025, doi: 10.1007/s00704-05912-2.

[2] 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 Agric. Technol., vol. 10, p. 100718, 2025, doi: 10.1016/j.atech.2024.100718.

[3] S. K. Roy and M. A. Uddin, “Prediction of rice yield in relation to weather variables in Bangladesh: Application of machine learning techniques,” Next Res., vol. 2, no. 3, p. 100653, 2025, doi: 10.1016/j.nexres.2025.100653.

[4] 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, 2024, doi: 10.3390/agriculture14091615.

[5] H. A. Salman, A. Kalakech, and A. Steiti, “Random forest algorithm overview,” Baghdad J. Mach. Learn., vol. 2024, pp. 69–79, 2024, doi: 10.58496/BJML/2024/007.

[6] B. Basso and L. Liu, “Seasonal crop yield forecast: Methods, applications, and accuracies,” in Advances in Agronomy, vol. 154, D. L. Sparks, Ed. Academic Press, 2019, pp. 201–255, doi: 10.1016/bs.agron.2018.11.002.

[7] A. Jabed, M. Azrifah, and A. 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, 2024, doi: 10.1016/j.heliyon.2024.e40836.
[8] S. Sah, D. Haldar, R. N. Singh, B. Das, and A. S. Nain, "Rice yield prediction through integration of biophysical parameters with SAR and optical remote sensing data using machine learning models," Sci. Rep., vol. 14, no. 1, p. 21674, Sep. 2024, doi: 10.1038/s41598-024-72624-4.
[9] M. Meroni, F. Waldner, L. Seguini, H. Kerdiles, and F. Rembold, "Yield forecasting with machine learning and small data: What gains for grains?" Agric. For. Meteorol., vol. 308–309, p. 108555, 2021, doi: 10.1016/j.agrformet.2021.108555.
[10] D. B. Lobell, W. Schlenker, and J. Costa-Roberts, "Climate trends and global crop production since 1980," Science, vol. 333, no. 6042, pp. 616–620, 2011, doi: 10.1126/science.1204531.
[11] H. Hersbach et al., "The ERA5 global reanalysis," Q. J. R. Meteorol. Soc., vol. 146, no. 730, pp. 1999–2049, 2020, doi: 10.1002/qj.3803.
[12] N. Hikmah, Asrirawan, and Apriyanto, "Peramalan data cuaca ekstrim Indonesia menggunakan model ARIMA dan recurrent neural network," Jambura J. Math., vol. 5, no. 1, pp. 230–242, 2023, doi: 10.34312/jjom.v5i1.17496.
[13] Statistics of Climate, Crop Pest and Climate Change Impact [Organisme Pengganggu Tanaman dan Dampak Perubahan Iklim]. Badan Pusat Statistik (BPS), Indonesia.
[14] P. Kumar, "An overview of Justicia adhatoda: A medicinal plant but native invader in India," Conservation, vol. 5, no. 1, p. 2, 2025, doi: 10.3390/conservation5010002.
[15] W. M. Isnaini and A. Aswi, "Metode geographically weighted LASSO dalam pemodelan tingkat pengangguran terbuka di Sulawesi Selatan," Econ. J. Stat. Its Appl., vol. 6, no. 1, pp. 70–78, 2025, doi: 10.20956/ejsa.v6i1.30863.
[16] L. A. Silva and G. Zanella, "Robust leave-one-out cross-validation for high-dimensional Bayesian models," J. Am. Stat. Assoc., vol. 119, no. 547, pp. 2369–2381, Jul. 2024, doi: 10.1080/01621459.2023.2257893.
[17] T. O. Hodson, "Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not," Geosci. Model Dev., vol. 15, no. 14, pp. 5481–5487, 2022, doi: 10.5194/gmd-15-5481-2022.
[18] R. Qiu, S. Xu, and Z. Yu, "Neural networks meet random forests," J. R. Stat. Soc. Ser. B Stat. Methodol., vol. 86, no. 5, pp. 1435–1454, 2024, doi: 10.1093/jrsssb/qkae038.
[19] S. Alimagham, M. P. van Loon, J. Ramirez-Villegas, H. N. C. Berghuijs, T. S. Rosenstock, and M. K. van Ittersum, "Integrating crop models and machine learning for projecting climate change impacts on crops in data-limited environments," Agric. Syst., vol. 228, p. 104367, 2025, doi: 10.1016/j.agsy.2025.104367.
[20] N. Roustaei, "Application and interpretation of linear-regression analysis," Med. Hypothesis Discov. Innov. Ophthalmol., vol. 13, no. 3, pp. 151–159, 2024, doi: 10.51329/mehdiophthal1506.
[21] D. Boldini, F. Grisoni, D. Kuhn, L. Friedrich, and S. A. Sieber, "Practical guidelines for the use of gradient boosting for molecular property prediction," J. Cheminform., vol. 15, no. 1, p. 73, 2023, doi: 10.1186/s13321-023-00743-7.
[22] P. S. Reddy, B. Sivanandini, M. Chaudhary, and N. Gautam, "Adaptive ML and DL framework for climate-resilient agriculture," Int. J. Innov. Res. Technol., vol. 10, no. 2, pp. 65–71, 2025.
[23] H. Huang, J. Liu, A. Chen, M. Ruiz-Vásquez, and R. Orth, "State-of-the-art hydrological datasets exhibit low water balance consistency globally," Earth Syst. Sci. Data Discuss., pp. 1–23, 2025, doi: 10.5194/essd-2025-376.
[24] S. E. El-Mahroug et al., "Predictive modeling of climate-driven crop yield variability using DSSAT towards sustainable agriculture," AgriEngineering, vol. 7, no. 5, p. 156, 2025, doi: 10.3390/agriengineering7050156.
[25] C. Han, S. Shi, J. Ji, and C. Peng, "Deep learning-based self-updating modeling approach for electro-hydrostatic actuators," in Proc. 8th Int. Conf. Intell. Robot. Control Eng. (IRCE), 2025, pp. 283–288, doi: 10.1109/IRCE66030.2025.11203082.