International Research Journal of Multidisciplinary Scope (IRJMS), 2026;7(3):1530-1546

Original Article | ISSN (O): 2582-631X DOI:10.47857/irjms.2026.v07i03.012859

Optimized Weighted Fusion for Rice Productivity Prediction Using Multi-source Agricultural Data

R Hadapiningradja Kusumodestoni[1,2]*, Fatchul Arifin[1], Mutiara Nugraheni[1]

[1]Doctoral Program in Engineering Science, Faculty of Engineering, Universitas Negeri Yogyakarta, Yogyakarta, Indonesia, [2]Faculty of Science and Technology, Universitas Islam Nahdlatul Ulama Jepara, Central Java, Indonesia. *Corresponding Author's Email: kusumodestoni@unisnu.ac.id

Abstract

Accurate rice productivity prediction is essential for agricultural planning, food-security monitoring and climate-risk mitigation. In this manuscript, rice productivity refers to the average harvested paddy yield per unit harvested area at the district/city-year level, expressed as quintal per hectare (q ha-1). This study proposes a validation-based optimized weighted fusion model that integrates Random Forest (RF), Extreme Gradient Boosting (XGBoost) and Long Short-Term Memory (LSTM) for rice productivity prediction using multi-source agricultural data from districts/cities in Central Java, Indonesia, during 2010-2025. The novelty of this approach lies in its constrained, non-negative weight optimization using a separate validation period, its controlled comparison against standalone, late-fusion, stacking and boosting-based alternatives and its robustness assessment under explicitly defined climate categories. Predictors included agricultural area, climate, soil-land, agronomic and engineered temporal rainfall features, while rice production was excluded to avoid target leakage. Data from 2010-2020 were used for training, 2021-2022 for validation and fusion-weight selection and 2023-2025 for independent future-period testing. The selected weights were 0.30 for RF, 0.60 for XGBoost and 0.10 for LSTM. The optimized fusion achieved the lowest RMSE (4.2847) and the highest R^2 (0.2812), while XGBoost remained slightly better in MAE and MAPE. Robustness analysis showed lower error under normal climate conditions and acceptable relative error under combined extreme conditions. SHAP analysis identified drainage type, soil type, year, sunshine duration, harvested area and rainfall lag features as key predictors. The framework improves squared-error stability and provides interpretable evidence for agricultural decision support.

Keywords: Climate Variability, Optimized Weighted Fusion, Rice Productivity Prediction, Xgboost.

Introduction

Rice productivity prediction is an essential component of agricultural decision support because it helps anticipate yield variability, food-supply stability and climate-related production risks in rice-producing regions. This issue is particularly important in Indonesia, where rice remains a strategic staple crop and where changes in productivity can influence food availability, farmer income, market supply and policy planning. At the district/city level, rice productivity is shaped by non-linear interactions among rainfall, temperature, humidity, sunshine duration, soil properties, drainage, irrigation reliability, pest incidence, planted area, harvested area and temporal rainfall memory. Recent food-security and climate reports emphasize that agricultural production systems are increasingly exposed to rainfall variability, drought, floods and other climate-related disruptions that may affect staple food availability (1, 2). Therefore, forecasting rice productivity requires models that can integrate agricultural, climate, soil-land, agronomic and temporal information in a statistically reliable and interpretable manner. Machine learning has been widely applied in crop-yield and productivity prediction because it can model complex relationships among heterogeneous variables that are difficult to capture using purely linear approaches. Previous reviews have reported that weather indicators, soil properties, crop-management factors, remote-sensing variables and spatial information can improve agricultural prediction when they are integrated with suitable algorithms (3-5). Random Forest (RF) is often used for agricultural tabular data because it handles non-linear interactions and reduces variance through ensemble decision trees (6). Extreme Gradient Boosting (XGBoost) is also widely adopted because it learns sequential residual corrections while controlling model complexity

(Received 11th May 2026; Accepted 15th July 2026; Published 25th July 2026)

through regularization (7). Long Short-Term Memory (LSTM) networks can represent sequential patterns in time-dependent variables (8). In addition, SHapley Additive exPlanations (SHAP) can translate model predictions into feature-level contributions, thereby improving model interpretation for decision makers (9).

Despite these developments, existing agricultural prediction studies still leave several method-logical gaps. Some studies report only the best standalone model without testing whether a hybrid or ensemble approach genuinely improves future-period generalization. Other studies use simple averaging, which may degrade prediction when a weak learner receives the same influence as a stronger learner. Stacking can also become unstable when the meta-learner is trained on limited prediction data rather than on a robust validation or out-of-fold procedure. In addition, several studies still use random data splitting for agricultural time series, which can inflate performance by mixing past and future observations. These limitations reduce confidence when models are expected to support official agricultural monitoring, early-warning platforms, or farmer-advisory services (10-12). A more reliable design should therefore preserve temporal order, avoid target leakage and select model combinations without using future test information. Another unresolved issue is the relationship between predictive accuracy and climate-condition robustness. Rice productivity is not only affected by normal seasonal variability, but also by climate extremes such as drought, excessive rainfall, delayed water availability, drainage stress and pest pressure. Previous studies have shown that machine-learning models can be sensitive to climate patterns and local management conditions (13, 14). However, many crop-productivity studies evaluate only overall test error without separately examining normal, wet-extreme, dry-extreme and combined-extreme conditions. As a result, it remains difficult to determine whether a model performs consistently across climate circumstances relevant for agricultural risk management. This is important because a model with good average performance may still produce large errors during extreme years, when reliable early-warning information is most needed.

The current research gap is therefore the need for a leakage-controlled, validation-based and interpretable fusion framework for rice productivity prediction using multi-source agricultural data. The novelty of this study is not merely the combination of RF, XGBoost and LSTM. Instead, the study proposes a constrained non-negative weighted-output fusion in which model weights are selected only from the validation period and then fixed for independent future-year testing. This differs from boosting, which relies on a single sequential tree mechanism; from simple ensembles, which assign equal weights; from stacking, which requires a second-stage learner; and from many hybrid approaches that report only aggregate accuracy. The proposed framework jointly evaluates predictive accuracy, climate-condition robustness and SHAP-based interpretation (15-17). By combining these elements, the study provides an analytical basis for understanding when and why the model performs well or poorly.

Central Java, Indonesia, was selected as the study area because it is a major rice-producing region with diverse agroecological and management conditions. Differences in rainfall, drainage, soil type, irrigation systems, land characteristics and pest occurrence can influence productivity patterns, making the region suitable for evaluating a multi-source prediction framework (18-21). The district/city-year unit of analysis is relevant for policy use because agricultural statistics and climate-risk responses are commonly organized at administrative levels. This aggregation requires careful interpretation because district-level productivity represents average harvested paddy yield rather than individual farm-level yield (22-27). The overarching aim of this study is to develop an interpretable optimized weighted fusion framework for rice productivity prediction at the district/city-year level in Central Java, Indonesia (28-31). The specific objectives are:

(a) to define and construct rice productivity as harvested paddy yield per hectare using multi-source agricultural data;
(b) to compare RF, XGBoost, LSTM, late fusion, stacking and optimized weighted fusion under the same temporal forecasting design;
(c) to evaluate model robustness under normal, extremely wet, extremely dry and combined extreme climate conditions; and

(d) to interpret the strongest productivity drivers using SHAP (32-36).

The expected contribution is a transparent decision-support model that can complement existing agricultural monitoring systems rather than replace official statistics or expert judgment (37, 38).

Methodology

Study Area and Dataset

The study used district/city-level agricultural and environmental data from Central Java Province, Indonesia, covering 2010-2025. The research site is in Central Java, which lies approximately between 5.67°S-8.50°S and 108.50°E-111.50°E, with an approximate provincial centroid at 7.15°S and 110.14°E. The final dataset contained 560 district/city-year observations. Rice productivity was defined as the average harvested paddy yield per unit harvested area for each district/city and year, expressed in quintal per hectare (q ha-1). It was calculated from official production and harvested-area statistics as total paddy production divided by harvested area and converted to q ha-1 where necessary. Therefore, rice productivity in this study represents average district/city-level productivity rather than individual farm-level yield. The 2025 records refer to the available institutional records at the time of dataset compilation; if final official 2025 statistics are not yet available at submission, the authors should label these records as provisional or restrict the final analysis period to confirmed records. Predictors included agricultural variables, climate variables, soil and land variables, agronomic variables and engineered temporal rainfall features. Agricultural variables included planted area and harvested area. Climate variables included rainfall, humidity, average temperature, sunshine duration, wind speed and extreme rainfall. Soil and land variables included soil type, soil pH, soil texture, soil moisture and drainage type. Agronomic variables included irrigation system and dominant pest type. Temporal features included rainfall lag-1, lag-2, lag-3, a three-period rainfall moving average, rainfall anomaly, SPI and extreme dry/wet flags.

Table 1 summarizes the data components used in this study. The table is placed in the Methodology section because it clarifies the source and role of each data group before model development. These data components were harmonized into a district/city-year unit of analysis before preprocessing and model training.

Table 1: Main Components of the Multi-source Dataset

Data Component	Variables or Examples	Data Source and Documentation to Cite
Agricultural data	Planted area, harvested area, productivity	BPS-Statistics Indonesia; BPS Province of Central Java; provincial/district agricultural agency records
Climate data	Rainfall, humidity, temperature, sunshine duration, wind speed, extreme rainfall	BMKG station data, official climate summaries, or equivalent gridded climate products
Soil and land data	Soil type, soil texture, soil pH, soil moisture, drainage type	Land-resource maps, soil databases and institutional land-information records
Agronomic data	Irrigation system, dominant pest type	Provincial/district agricultural agency records and crop-protection reports
Temporal features	Rain_Lag_1, Rain_Lag_2, Rain_Lag_3, Rain_MA_3, SPI, Rain_Anomaly	Engineered by the authors from harmonized rainfall series

Note: BPS = Statistics Indonesia; BMKG = Meteorology, Climatology and Geophysical Agency; MA = Moving Average; Rain_MA_3 = three-period rainfall moving average; SPI = Standardized Precipitation Index.

Data Source and Unit of Analysis

The unit of analysis was district/city-year. Agricultural statistics were compiled from BPS-Statistics Indonesia, BPS Province of Central Java and provincial or district agricultural agency records, including harvested area, planted area, production and productivity records derived from official survey, administrative and Area Sampling Frame-based reporting systems where available. Climate variables were compiled from meteorological records provided by the Meteorology, Climatology and Geophysical Agency (BMKG) or equivalent station/gridded climate products. Daily or monthly climate observations were aggregated into annual district/city indicators, including rainfall totals or averages, humidity, average temperature, sunshine duration, wind speed, rainfall anomaly and SPI. Soil-land variables were compiled from land-resource maps and institutional soil information, while agronomic variables such as irrigation system and dominant pest type were obtained from agricultural agency records. The spatial coverage included districts/ cities in Central Java with complete matched records across agricultural, climate, soil-land and agronomic sources during the study period. Because the original sources used different spatial and temporal resolutions, all variables were harmonized to district/city-year observations before modeling by matching administrative codes, year labels, units and variable definitions. Quality assurance included duplicate-record checks, unit consistency checks, missing-value inspection, range and outlier screening, cross-comparison with official annual trends and verification that production was excluded from the predictor set to avoid target leakage.

Preprocessing and Leakage Control

Preprocessing included checking data structure, missing values, variable types, duplicate records, anomalous entries and consistency between district/city identifiers and observation years. The final modeling dataset did not contain missing values. Categorical variables such as region, soil type, soil texture, soil moisture, drainage type, irrigation system and dominant pest type were encoded into numerical codes. Quality assurance was performed before modeling by standardizing measurement units, comparing descriptive statistics with official summaries, identifying impossible or implausible values and correcting only clearly documented data-entry anomalies. A constant fertilizer-code variable was excluded because it had no useful variation. An anomalously low average-temperature value was corrected using the median to reduce the influence of data-entry error. Rice production was removed from the feature set because rice productivity is mathematically derived from production and harvested area. This leakage-control step ensured that models learned from explanatory agricultural, climate, soil and temporal predictors rather than from direct target information.

Climate-aware Feature Engineering

Feature engineering was conducted to represent delayed and accumulated climate effects. Rain_Lag_1, Rain_Lag_2 and Rain_Lag_3 represented rainfall from previous periods. Rain_MA_3 represented a three-period moving average of rainfall. Rain_Anomaly represented the deviation of rainfall from the historical mean. SPI was calculated by standardizing rainfall using the historical mean and standard deviation. Extreme_Dry_Flag and Extreme_Wet_Flag identified dry and wet extreme observations. These features were included because rice productivity may respond not only to current rainfall, but also to previous water availability, soil moisture carry-over, planting decisions, irrigation reliability and pest dynamics.

Temporal Train-validation-test Design

To simulate a realistic forecasting scenario, the dataset was split temporally into three non-overlapping segments. Observations from 2010-2020 were used for base-model training, observations from 2021-2022 were used for validation and fusion-weight selection and observations from 2023-2025 were retained as an independent future-period test set. This division was considered sufficient for the present district/city-year forecasting task because the 2010-2020 training period covers eleven annual rice-production cycles and therefore captures interannual climatic variability, recurrent wet and dry phases and long-term district-level agricultural patterns. The 2021-2022 validation period provides a chronologically later tuning window for selecting hyperparameters and fusion weights without using future test information. The 2023-2025 test period then evaluates whether the final model generalizes to the most recent unseen years, which is closer to a real forecasting application than random splitting. Although this split cannot represent every possible long-term climate regime, it preserves temporal ordering, reduces optimistic bias and is complemented by a separate robustness analysis under normal, extremely wet, extremely dry and combined extreme climate conditions. Hyperparameters and fusion weights were not selected using the test set.

Base Learners and Hyperparameters

RF was used as a variance-reduction learner for non-linear tabular patterns. XGBoost was used as a regularized boosting learner for strong structured-data prediction. LSTM was used as a temporal learner to capture sequential information from lagged rainfall, SPI, rainfall anomaly and other climate-related features. The hyperparameter optimization process was conducted in a validation-based manner. For RF, candidate settings varied the number of trees, maximum depth, minimum split size, minimum leaf size and feature-sampling strategy; the final configuration was selected according to validation RMSE while maintaining stable out-of-sample error. For XGBoost, the tuning process examined learning rate, number of estimators, maximum depth, subsampling ratio, column-sampling ratio and L2 regularization, with early stopping monitored on the validation period where applicable. For LSTM, sequence length, hidden units, dropout rate, batch size, learning rate, number of epochs and early-stopping patience were tuned using validation RMSE and loss convergence. Numerical predictors were scaled where required by the algorithm, categorical predictors were encoded consistently within the temporal split and random seeds were fixed to improve reproducibility. The final implementation settings are summarized in Table 2 and Table 3.

Table 2 describes the functional role of each model in the proposed framework. RF provides stable tabular learning, XGBoost serves as the dominant structured-data learner and LSTM captures temporal climate representations. The table also explains that optimized weighted fusion integrates the prediction outputs rather than raw variables.

Table 2: Model Roles in the Proposed Framework

Model	Main Role	Input Emphasis
Random Forest	Variance reduction and tabular pattern learning	Soil, area, climate and agronomic variables
XGBoost	Residual correction and strong tabular prediction	All encoded tabular and engineered features
LSTM	Sequential climate pattern learning	Lagged rainfall, SPI, rainfall anomaly and climate variables
Optimized weighted fusion	Validation-based integration of model predictions	Pred_RF, Pred_XGB and Pred_LSTM

Note: RF = Random Forest; XGBoost = Extreme Gradient Boosting; XGB = XGBoost; LSTM = Long Short-Term Memory; SPI = Standardized Precipitation Index; Pred = prediction or predicted output; Pred_RF = prediction generated by Random Forest; Pred_XGB = prediction generated by XGBoost; Pred_LSTM = prediction generated by Long Short-Term Memory.

Table 3 provides the final hyperparameter configuration used for reporting. This table improves reproducibility because it allows other researchers to rerun the base learners and fusion process under the same settings. The validation criterion and selected weights are included to clarify that the test set was not used during model selection.

Table 3: Final Model Configuration for Reproducibility

Model	Final Configuration Used for Reporting
Random Forest	n_estimators = 500; max_depth = None; min_samples_split = 2; min_samples_leaf = 1; max_features = sqrt; bootstrap = True; random_state = 42. Purpose: robust tabular baseline and complementary learner.
XGBoost	objective = reg:squarederror; n_estimators = 300; max_depth = 3; learning_rate = 0.05; subsample = 0.80; colsample_bytree = 0.80; reg_lambda = 1.0; random_state = 42. Purpose: dominant structured-data learner.
LSTM	sequence_length = 3; LSTM units = 32; dropout = 0.20; dense output = 1; optimizer = Adam; learning_rate = 0.001; loss = MSE; epochs = 200; batch_size = 16; early_stopping_patience = 20; scaler = MinMaxScaler. Purpose: temporal learner for lagged climate signals.
Optimized fusion	validation criterion = RMSE; non-negative weights; sum of weights = 1; grid step = 0.05; selected weights: RF = 0.30, XGBoost = 0.60, LSTM = 0.10. Purpose: validation-based prediction integration.

Note: RF = Random Forest; XGBoost = Extreme Gradient Boosting; LSTM = Long Short-Term Memory; RMSE = root mean squared error; MSE = mean squared error; sqrt = square root; reg:squarederror = regression with squared-error objective; MinMaxScaler = minimum–maximum feature scaling method.

Optimized Weighted Fusion

The final weighted prediction was computed using Equation [1]. This equation is cited here because it defines the fusion output before the complete mathematical formulation is provided. Candidate weights were searched on the validation set using RMSE as the primary criterion with a grid step of 0.05 and a non-negative weight constraint. After the best weights were selected, they were fixed and evaluated on the independent future test set.

$$\hat{y}_i = w_{RF} \cdot y_{RF,i} + w_{XGB} \cdot y_{XGB,i} + w_{LSTM} \cdot y_{LSTM,i}, \quad w_{RF} + w_{XGB} + w_{LSTM} = 1 \quad [1]$$

The selected combination was w_RF = 0.30, w_XGB = 0.60 and w_LSTM = 0.10, indicating that XGBoost contributed the strongest predictive signal, RF provided complementary tabular stability and LSTM contributed a smaller temporal signal.

Evaluation Metrics and Interpretability

Model performance was evaluated using root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and coefficient of determination (R^2). RMSE was used as the primary selection criterion because it penalizes larger errors. MAE and MAPE were used to assess absolute and relative prediction error. R^2 was used to evaluate the proportion of target variation explained by the model. Climate-condition classes were defined using SPI and rainfall-anomaly thresholds to ensure reproduce-bility. Normal climate observations were defined as -1.0 < SPI < 1.0 and rainfall anomaly between -1 and +1 standard deviation. Extremely dry observations were defined as SPI <= -1.0 or rainfall anomaly <= -1 standard deviation. Extremely wet observations were defined as SPI >= 1.0 or rainfall anomaly >= +1 standard deviation. Combined extreme observations were defined as the union of extremely dry and extremely wet observations. These thresholds follow the common interpretation that SPI values between -1.0 and +1.0 indicate near-normal precipitation, while values beyond these limits represent moderate to severe wet or dry departures. Robustness was then assessed by calculating the same error metrics within each climate subset. SHAP analysis was applied to the XGBoost component because it received the largest fusion weight and represented the dominant tabular learner in the final ensemble. The mathematical expressions of the performance metrics and model outputs are provided in the mathematical formulation subsection.

Figure 1 illustrates the complete modeling workflow used in this study. The workflow starts with multi-source agricultural data and applies preprocessing and leakage control before training RF, XGBoost and LSTM. The model outputs are combined using validation-selected weights and the final predictions are evaluated on the independent future test period. The figure also shows that robustness analysis and SHAP interpretation are performed after test-year prediction.

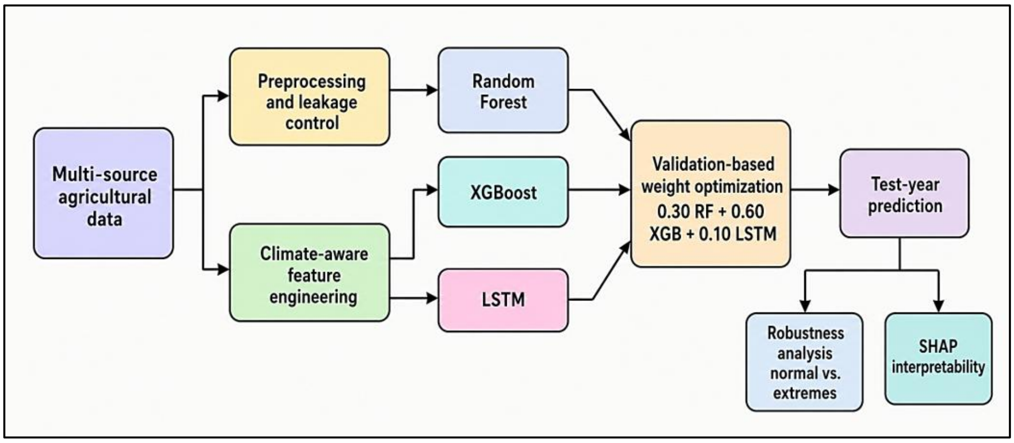

Figure 1: Experimental Workflow for Validation-based Optimized Weighted Fusion. The Workflow Shows Data Preprocessing, Climate-aware Feature Engineering, Base Model Training, Validation-based Weight Optimization, Future-year Testing, Robustness Analysis and SHAP Interpretation

Figure 2 presents the temporal trend of average rice productivity from 2010 to 2025. The graph helps contextualize the train-validation-test split by showing the variation observed across years. The visible fluctuations indicate that the dataset contains both productivity increases and decreases, which is relevant for evaluating future-period forecasting ability. The error bars summarize the uncertainty of district/city observations within each year.

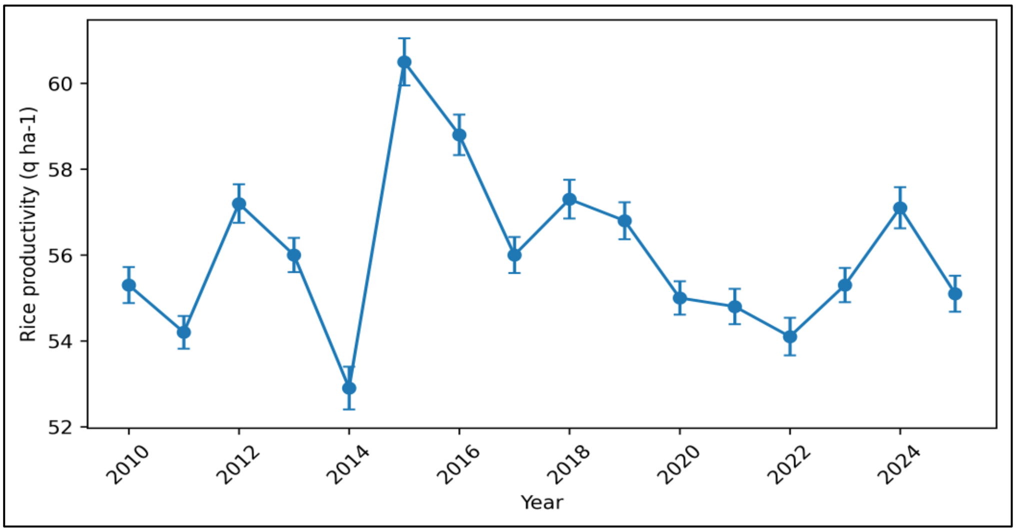

Figure 2: Average Rice Productivity Trend in the Study Dataset. Error Bars Indicate the Standard Error Across District/City Observations and Grid Lines Were Removed for Journal Artwork Compliance. The X Axis Represents Observation Year and the Y Axis Represents Rice Productivity in Q Ha-1

Mathematical Formulation and Model Output Equations

To strengthen reproducibility, this subsection presents the equations that form the basis and output of the target construction, climate-aware feature engineering, base learners, fusion strategies, evaluation metrics, robustness analysis and SHAP interpretation.

Let x_i denote the predictor vector of observation i, y_i denote the observed productivity value and y_hat_i denote the predicted productivity value.

Rice productivity is expressed as harvested paddy yield per hectare at the district/city-year level, as shown in Equation [2]. Production was not used as an input predictor because productivity is mathematically derived from production and harvested area.

$$y_i = \mathrm{Productivity}_i = \frac{\mathrm{Production}_i}{\mathrm{HarvestedArea}_i} \qquad [2]$$

Climate-aware temporal features were constructed to represent rainfall memory, accumulated rainfall, rainfall anomaly and standardized precipitation Equations [3-6].

$$\mathrm{RainLag}_{k,t} = \mathrm{Rain}_{t-k}, \quad k \in \{1,2,3\} \qquad [3]$$

$$\mathrm{RainMA}_{3,t} = \frac{\mathrm{Rain}_t + \mathrm{Rain}_{t-1} + \mathrm{Rain}_{t-2}}{3} \qquad [4]$$

$$\mathrm{RainAnomaly}_t = \mathrm{Rain}_t - \mu_{rain} \qquad [5]$$

$$\mathrm{SPI}_t = \frac{\mathrm{Rain}_t - \mu_{rain}}{\sigma_{rain}} \qquad [6]$$

The Random Forest prediction is obtained by averaging the outputs of B regression trees, as expressed in Equation [7]:

$$\hat{y}_{RF}(x_i) = \frac{1}{B}\sum_{b=1}^{B} T_b\,(x_i) \qquad [7]$$

The XGBoost prediction is an additive ensemble of M regression trees. Equations [8-10] describe the additive output, objective function and regularized loss.

$$\hat{y}_{XGB}(x_i) = \sum_{m=1}^{M} f_m\,(x_i), \quad f_m \in \mathcal{F} \qquad [8]$$

$$\mathrm{Obj}^{(m)} = \sum_{i=1}^{n} \ell\left(y_i, \hat{y}_i^{(m-1)} + f_m(x_i)\right) + \Omega(f_m) \qquad [9]$$

$$\Omega(f) = \gamma K + \frac{1}{2}\lambda \sum_{j=1}^{K} w_j^2 \quad [10]$$

The LSTM learner models temporal climate representations through the gate, cell-state and hidden-state equations shown in Equations [11-17].

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad [11]$$

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad [12]$$

$$\tilde{c}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad [13]$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad [14]$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad [15]$$

$$h_t = o_t \odot \tanh(c_t) \quad [16]$$

$$\hat{y}_{LSTM} = W_y h_T + b_y \quad [17]$$

Late-fusion baselines used unweighted averaging, while the optimized weighted fusion used validation-selected non-negative weights. These outputs are formalized in Equations [18-22].

$$\hat{y}_{LF-RF-XGB} = \frac{\hat{y}_{RF} + \hat{y}_{XGB}}{2} \quad [18]$$

$$\hat{y}_{LF-RF-XGB-LSTM} = \frac{\hat{y}_{RF} + \hat{y}_{XGB} + \hat{y}_{LSTM}}{3} \quad [19]$$

$$\hat{y}_{OWF} = w_{RF}\hat{y}_{RF} + w_{XGB}\hat{y}_{XGB} + w_{LSTM}\hat{y}_{LSTM} \quad [20]$$

$$w_{RF} + w_{XGB} + w_{LSTM} = 1, \quad w_j \geq 0 \quad [21]$$

$$\left(w_{RF}^*, w_{XGB}^*, w_{LSTM}^*\right) = \arg\min_{w} \ RMSE_{val}(y, \hat{y}_{OWF}) \quad [22]$$

The stacking baseline uses the predictions of the base models as meta-features and learns a second-stage mapping, as shown in Equations [23, 24].

$$z_i = [\hat{y}_{RF}(x_i), \hat{y}_{XGB}(x_i)] \quad [23]$$

$$\hat{y}_{STACK}(x_i) = g(z_i) \quad [24]$$

The evaluation outputs were calculated using RMSE, MAE, MAPE and R2 in Equations [25-28]. Robustness analysis and SHAP interpretation are described in Equations [29, 30].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad [25]$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad [26]$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad [27]$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad [28]$$

$$RMSE_c = \sqrt{\frac{1}{|D_c|}\sum_{i \in D_c}(y_i - \hat{y}_i)^2} \quad [29]$$

$$f(x_i) = \phi_0 + \sum_{j=1}^{p}\phi_j(x_i) \quad [30]$$

Results

Baseline Performance

The baseline results show that XGBoost was the strongest standalone model. RF achieved RMSE = 4.3690, MAE = 3.3549, MAPE = 5.9287% and R^2 = 0.2526. XGBoost improved performance with RMSE = 4.3244, MAE = 3.1605, MAPE = 5.5069% and R^2 = 0.2678. LSTM produced weaker standalone results, with RMSE = 8.0274 and R^2 = -0.6592. This indicates that the annual-resolution sequence representation was insufficient for LSTM to outperform tree-based tabular models. The finding is important because it shows that model complexity alone does not guarantee improved agricultural prediction.

Table 4 presents the baseline performance of the standalone models. The table is important because it establishes the reference point before evaluating fusion alternatives. XGBoost provides the strongest standalone result, whereas LSTM performs poorly in this annual-resolution setting. This confirms that all fusion results should be interpreted relative to a strong XGBoost baseline.

Table 4: Baseline Model Performance

Model	RMSE	MAE	MAPE (%)	R2
Random Forest	4.369046	3.354935	5.928732	0.252605
XGBoost	4.324424	3.160500	5.506889	0.267794
LSTM	8.027424	5.921450	9.488234	-0.659247

Note: RF = Random Forest; XGBoost = Extreme Gradient Boosting; LSTM = Long Short-Term Memory; RMSE = root mean squared error; MAE = mean absolute error; MAPE = mean absolute percentage error; R2 = coefficient of determination. Lower RMSE, MAE and MAPE and higher R2 indicate better performance.

Figure 3 compares RMSE values across standalone and fusion models. The optimized weighted fusion achieved the lowest RMSE, while the LSTM and stacking baselines produced larger errors. The figure provides visual evidence that validation-based weighting can improve squared-error performance over naive model combination. The error bars indicate uncertainty in test residuals, which helps evaluate the stability of the differences.

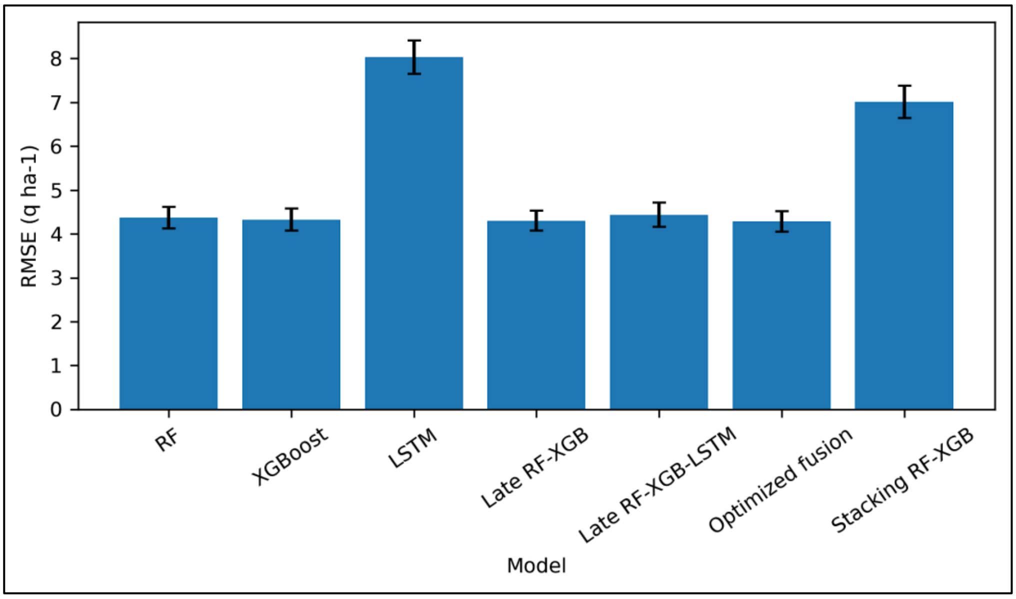

Figure 3: RMSE Comparison Across Evaluated Models. Error Bars Represent Bootstrap-Based Uncertainty Intervals from Test Residuals; Lower Values Indicate Better Squared-error Performance

Figure 4 presents the relative prediction error across all evaluated models. XGBoost remains slightly better in MAPE than the optimized fusion, even though the optimized fusion has the lowest RMSE. This difference is important because it shows that model ranking can vary across metrics. The figure supports the interpretation that the proposed method primarily improves squared-error stability rather than every error criterion.

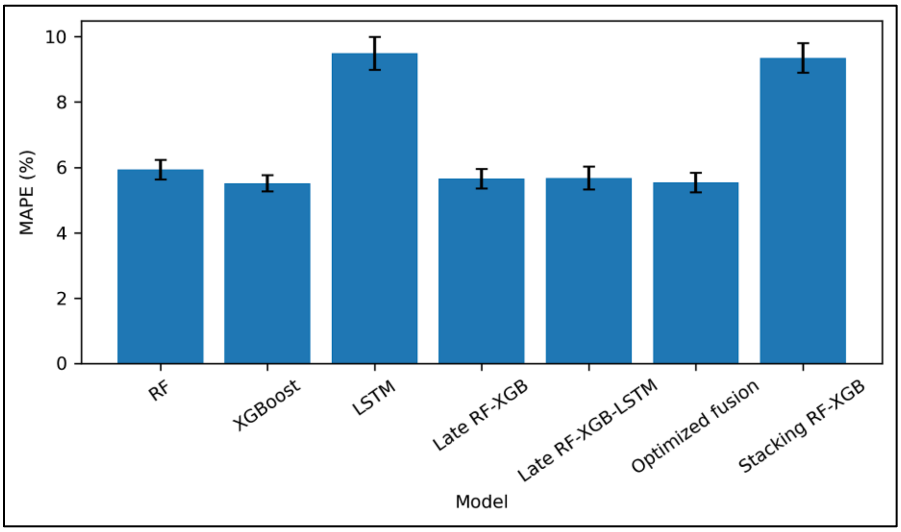

Figure 4: MAPE Comparison Across Evaluated Models. Error Bars Represent Bootstrap-based Uncertainty Intervals from Test Residuals; Lower Values Indicate Better Relative Error Performance

Fusion Performance

The fusion experiments demonstrate that the choice of fusion strategy strongly affects prediction quality. Late Fusion RF-XGB achieved RMSE = 4.3002 and R^2 = 0.2760, improving slightly over the standalone tree-based models in RMSE. However, Late Fusion RF-XGB-LSTM achieved RMSE = 4.4366, indicating that assigning equal influence to a weak LSTM component can degrade combined prediction. Stacking RF-XGB performed poorly in this preliminary setting, with RMSE = 7.0135, suggesting that stacking requires proper out-of-fold construction and sufficient meta-training data. In contrast, optimized weighted fusion achieved the best RMSE (4.2847) and R^2 (0.2812), confirming that fusion is beneficial when learner contributions are selected through validation rather than assumed equal.

Table 5 extends the baseline comparison by including late fusion, optimized weighted fusion and stacking. It shows that not all fusion methods improve performance. Equal-weight fusion becomes less effective when the weaker LSTM component is included. The optimized fusion provides the most balanced improvement by limiting the LSTM weight and retaining the dominant contribution of XGBoost.

Table 5: Performance Comparison of Baseline and Fusion Models

Model	RMSE	MAE	MAPE (%)	R2
Random Forest	4.369046	3.354935	5.928732	0.252605
XGBoost	4.324424	3.160500	5.506889	0.267794
LSTM	8.027424	5.921450	9.488234	-0.659247
Late Fusion RF-XGB	4.300160	3.223313	5.656415	0.275987
Late Fusion RF-XGB-LSTM	4.436620	3.284496	5.668422	0.229307
Optimized Weighted Fusion RF-XGB-LSTM	4.284739	3.177932	5.535749	0.281171
Stacking RF-XGB	7.013492	5.703519	9.353674	0.077371

Note: RF = Random Forest; XGBoost = Extreme Gradient Boosting; LSTM = Long Short-Term Memory. RMSE, MAE, MAPE and R2 are calculated on the independent 2023-2025 test period. The optimized fusion used weights selected only from the validation period.

Robustness Under Climate Conditions

Robustness analysis was conducted by separating test observations into normal, extremely wet, extremely dry and combined extreme climate conditions using the SPI and rainfall-anomaly thresholds defined in the Methodology section. Normal observations satisfied -1.0 < SPI < 1.0 and rainfall anomaly within +/-1 standard deviation; extremely dry observations satisfied SPI <= -1.0 or rainfall anomaly <= -1 standard deviation; extremely wet observations satisfied SPI >= 1.0 or rainfall anomaly >= +1 standard deviation; and combined extreme observations consisted of both extremely dry and extremely wet cases. The optimized weighted fusion model performed best under normal conditions, achieving RMSE = 3.4029 and MAPE = 4.5998%. Under combined extreme conditions, RMSE increased to 6.4218 and MAPE increased to 8.6945%. Although error increased during climate extremes, the relative error remained below 10%, suggesting acceptable but not perfect robustness.

Table 6 evaluates the final model under different climate conditions. This table is essential because a model with acceptable overall performance may still fail during extreme climate periods. The result shows that error increases under wet and dry extremes, but the relative error remains below 10%. This supports the use of the model as a decision-support tool with caution during climate anomalies.

Table 6: Robustness Analysis by Climate Condition

Climate Condition	N	RMSE	MAE	MAPE (%)	R2
Normal	81	3.402870	2.592194	4.599826	0.288200
Extremely wet	15	6.443048	4.918952	8.349985	-0.418123
Extremely dry	9	6.386261	5.547875	9.268670	0.593115
Combined extreme	24	6.421812	5.154798	8.694491	0.267072

Note: Normal = -1.0 < SPI < 1.0 and rainfall anomaly within +/-1 standard deviation; extremely dry = SPI <= -1.0 or rainfall anomaly <= -1 standard deviation; extremely wet = SPI >= 1.0 or rainfall anomaly >= +1 standard deviation; combined extreme = union of extremely dry and extremely wet cases.

Figure 5 visualizes the robustness results by comparing RMSE under normal and extreme climate conditions. Normal conditions produce the lowest RMSE, whereas extremely wet, extremely dry and combined extreme conditions generate higher errors. The figure confirms that climate extremes remain more difficult to predict. This visualization complements Table 6 by showing the magnitude of error differences across climate categories.

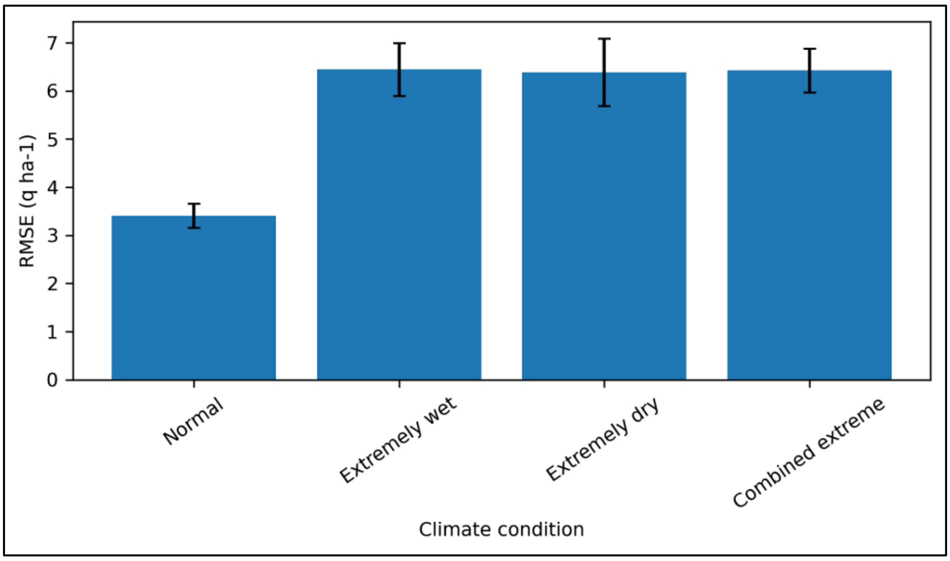

Figure 5: RMSE Under Normal and Extreme Climate Conditions. Error Bars Represent Uncertainty Intervals Across Test Observations Within Each Climate Subset

Interpretability Analysis

SHAP analysis was applied to the XGBoost component because it had the largest fusion weight. The most influential predictors were drainage type, soil type, year, sunshine duration, harvested area, Rain_Lag_1, humidity, Rain_Lag_3, soil pH and wind speed. These results indicate that rice productivity was not governed by climate alone. Land and water-management conditions played a central role in the prediction process. Drainage type had the highest mean absolute SHAP value, suggesting that water regulation and land hydrology were crucial for productivity. Soil type was also highly influential, reflecting the importance of physical soil conditions for water retention, nutrient availability and crop response. Rainfall lag features demonstrate that previous rainfall conditions contributed to productivity outcomes.

Table 7 ranks the strongest explanatory features derived from SHAP analysis. The table indicates that drainage and soil characteristics dominate the prediction process. Rainfall lags remain relevant but are not the only drivers of productivity. This reinforces the need to combine climate, soil-land and management-related variables in rice-productivity prediction.

Table 7. Top explanatory features based on SHAP

Rank	Feature	Mean |SHAP|
1	Drainage type	1.230348
2	Soil type	0.833692
3	Year	0.770592
4	Sunshine duration	0.698169
5	Harvested area	0.560543
6	Rain lag-1	0.397453
7	Humidity	0.380367
8	Rain lag-3	0.355690
9	Soil pH	0.346438
10	Wind speed	0.314964

Note: Mean |SHAP| indicates the mean absolute contribution of each predictor to the XGBoost component across test observations

Figure 6 presents the same SHAP ranking as a horizontal bar chart. The Y axis title has been added to clearly identify the feature dimension, while the X axis represents mean absolute SHAP value. Drainage type has the largest contribution, followed by soil type, year and sunshine duration. The figure makes the interpretability results easier to compare visually across predictors.

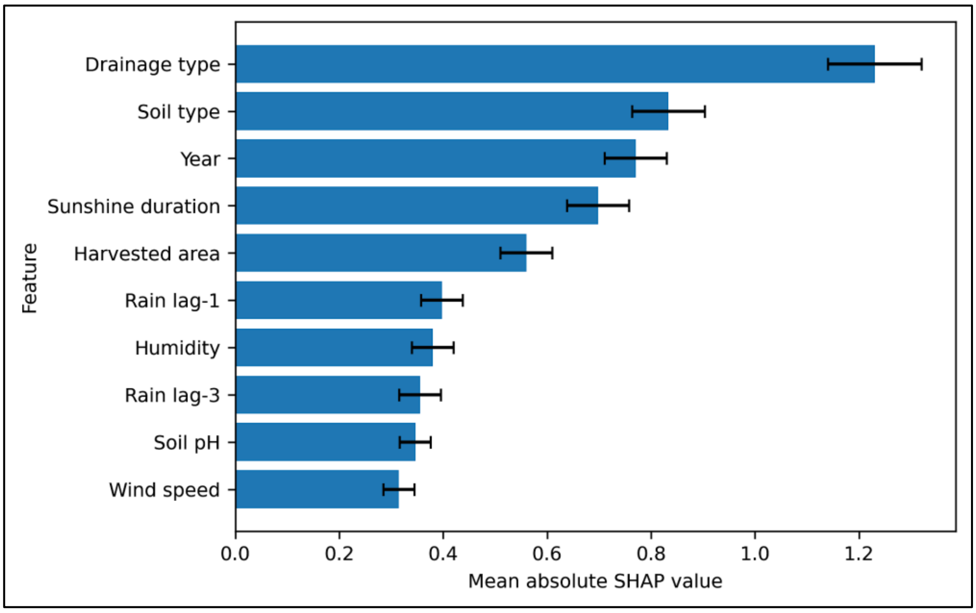

Figure 6: Top Ten Explanatory Features Based on SHAP Values. Error Bars Represent Variation in Absolute SHAP Values Across Test Observations

Discussion

The results support the argument that optimized fusion is more reliable than naive fusion. While XGBoost was the best single model, the optimized weighted fusion reduced RMSE and increased R^2. The improvement was modest, but meaningful because it was obtained under a future-period evaluation design and after comparing several baseline and fusion alternatives. The result also suggests that XGBoost should remain the dominant component in structured agricultural prediction tasks, while RF and LSTM can provide complementary information when their influence is properly controlled.

Comparative positioning with previous studies. The finding that XGBoost was the strongest standalone model is consistent with comparative crop-yield studies reporting that boosting algorithms are often competitive on structured agricultural datasets with mixed climate, soil and management predictors (23, 24, 28). However, the present study differs from those studies by showing that the best standalone boosting model can still be marginally improved in squared-error performance when its output is combined with RF and a small LSTM contribution through validation-selected weights. The weak standalone LSTM result is also consistent with studies noting that deep sequential models generally require dense monthly, seasonal, or remote-sensing time series to outperform tabular learners (7, 18, 35). Compared with recent hybrid and explainable agricultural models, this study contributes a simpler but more transparent fusion equation, explicit leakage control and climate-condition robustness analysis, thereby complementing studies that emphasize model complexity or image-based crop diagnosis (13, 20, 36, 37). These comparisons indicate that the proposed method is most useful when the available dataset is multi-source but still primarily tabular and annual, where transparent weighting may be more reliable than highly complex deep architectures.

Comparison with official monitoring and forecasting systems. In Indonesia, rice statistics and production outlooks are commonly supported by official agricultural statistics, administrative reporting, BMKG climate monitoring and Area Sampling Frame-based harvested-area estimation used by BPS. At the global level, monitoring programs such as the FAO Global Information and Early Warning System (GIEWS) and the Agricultural Market Information System (AMIS) combine official statistics, market information, climate indicators and expert assessments to track food supply conditions and emerging risks (35-37). The proposed framework differs from these operational systems because it focuses on district/city-level productivity prediction using machine-learning fusion and interpretable feature attribution.

Critical Interpretation and Limitations

The R^2 value of the final model remained relatively low, indicating that a substantial proportion of productivity variation was not explained by the available predictors. This should not be interpreted as model failure, but as evidence that rice productivity is influenced by additional variables that were not fully represented. Important omitted variables include labor availability, mechanization level, irrigation investment, agricultural subsidies, rice variety, seed quality, fertilizer dose and timing, planting calendar, pest intensity, irrigation volume, farming technology, extension support and satellite-derived vegetation indices. The absence of labor, mechanization, investment and subsidy variables may explain part of the unexplained variance because these factors influence the timeliness of cultivation, water-control capacity, input accessibility and farmers' ability to respond to climate stress. Therefore, the proposed model is best interpreted as an early-stage decision-support framework rather than a final operational forecasting system. In addition, the RMSE improvement over XGBoost was modest; therefore, the study emphasizes squared-error stability and interpretability rather than claiming overwhelming superiority. Future improvements should include higher-resolution temporal data, spatial-temporal validation, remote sensing indicators, richer agronomic and socioeconomic variables, out-of-fold stacking and bootstrap or paired-error statistical tests to quantify whether performance differences are statistically significant.

Applicability to Other Rice-growing Locations

The proposed framework is applicable to other rice-growing locations, but it should not be transferred without local recalibration. For irrigated lowland areas, the same framework can emphasize irrigation reliability, drainage, soil texture, planting area and rainfall-memory features. For rainfed, coastal, upland, or flood-prone rice systems, additional local variables such as tidal influence, groundwater availability, salinity, planting calendar, cultivar type, fertilizer management, pest intensity, labor availability, mechanization, irrigation investment, subsidy access and satellite-derived vegetation indices should be incorporated. For practical deployment, the model can be linked to government agricultural information systems by converting the trained pipeline into a dashboard or application programming interface that receives periodic BPS or agricultural-agency productivity records, BMKG rainfall and climate updates and local agronomic reports. District-level predicted productivity and SHAP-based risk drivers can then be displayed as early-warning indicators for agricultural offices, irrigation managers, extension workers and farmer-advising services. The outputs may support decisions on drainage improvement, irrigation prioritization, pest surveillance, planting-calendar adjustment and targeted advisory messages. The RF, XGBoost and LSTM components can be retained, but the optimal weights should be re-estimated using a validation period from the target location because the relative importance of climate, land and management factors may differ across agroecological zones. Therefore, the framework is best understood as a transferable modeling procedure rather than a fixed model with universal weights.

Conclusion

This study proposed and evaluated an optimized weighted fusion model integrating RF, XGBoost and LSTM for rice productivity prediction using multi-source agricultural data in Central Java, Indonesia. Rice productivity was defined as district/city-level average harvested paddy yield per hectare, expressed in q ha-1. The research used a temporal forecasting design in which data from 2010-2020 were used for model training, 2021-2022 for validation and fusion-weight selection and 2023-2025 for independent future-period evaluation. This design reduces the possibility of overly optimistic performance that may occur when past and future observations are randomly mixed.

The main contribution of the study is the development of a validation-based and leakage-controlled weighted-output fusion framework. The final fusion assigned weights of 0.30 to RF, 0.60 to XGBoost and 0.10 to LSTM. It achieved the lowest RMSE and the highest R2 among the tested models, although XGBoost remained slightly better in MAE and MAPE. These results imply that optimized fusion can improve squared-error stability, but the improvement should be interpreted carefully because the gain over the strongest standalone model was modest. The study also shows that equal weighting can be harmful when one component model performs poorly, as demonstrated by the inclusion of the weak standalone LSTM model in naive late fusion.

The robustness and interpretability findings provide additional practical consequences. The model performed better under normal climate conditions than under combined extreme conditions, indicating that climate extremes remain challenging for rice-productivity prediction. SHAP analysis identified drainage type, soil type, year, sunshine duration, harvested area and rainfall lag features as influential predictors. These findings suggest that productivity forecasting should not rely only on climate indicators; land, water-management, soil and agronomic factors are also essential for decision support.

The study has several limitations. The moderate R2 value indicates that a large proportion of productivity variation remains unexplained. The annual district/city-level structure may mask farm-level differences, seasonal crop dynamics and local management practices. Important variables such as labor availability, mechanization, irrigation investment, agricultural subsidies, cultivar choice, fertilizer management and satellite-derived vegetation indicators were not fully represented. In addition, the model was developed for Central Java and should be recalibrated before application to other rice-growing regions.

Future research should incorporate higher-resolution temporal data, spatial-temporal validation, remote-sensing indicators, more detailed agronomic and socioeconomic variables and paired-error statistical tests to evaluate whether differences among models are statistically significant. Further work should also explore operational dashboard design, integration with government agricultural information systems and farmer-advisory services. These developments would strengthen the practical value of the proposed framework while maintaining transparency, reproducibility and interpretability.

Abbreviations

MAPE: mean absolute percentage error, SHAP: SHapley Additive exPlanations, SPI: Standardized Precipitation Index, XAI: explainable artificial intelligence.

Acknowledgement

The authors thank Universitas Negeri Yogyakarta and the institutions providing agricultural and climate data used in this research. The authors also acknowledge the academic supervision and methodological feedback received during the dissertation development process.

Author Contributions

Hadapiningradja Kusumodestoni R: Conceptualization, Data Curation, Methodology, Analysis, Visualization, Writing- Original Draft Preparation, Fatchul Arifin: Supervision, Methodology Validation, Manuscript Review, Mutiara Nugraheni: Supervision, Interpretation, Manuscript Review. All authors read and approved the final manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability

The processed dataset, data dictionary, quality-assurance log, model-configuration file and Python scripts used in this study are available from the corresponding author upon reasonable request. Public release of the complete dataset may depend on data-sharing permissions from the original data providers. The data dictionary should include rice-productivity units, spatial identifiers, temporal aggregation rules, source documents and variable definitions to improve reproducibility.

Declaration of Artificial Intelligence (AI) Assistance

The authors used AI-assisted language support to improve grammar, formatting consistency and editorial clarity during revision. No AI tool was used to create the research data, conduct the statistical analysis, or make conceptual decisions. All data analysis, interpretation and final manuscript responsibility remain with the authors.

Ethics Approval

This study used secondary agricultural, climate, soil and agronomic data and did not involve human participants, patient information, or animal experiments. Therefore, human or animal ethics approval was not applicable.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. World Meteorological Organization. State of the Global Climate 2024. Geneva: World Meteorological Organization; 2025. WMO-No. 1368. ISBN: 978-92-63-11368-5. https://wmo.int/sites/default/files/2025-03/WMO-1368-2024_en.pdf
2. Bernard BM Jr, Song Y, Sumiati, Wang X. Assessing the impact of climate variability on total factor productivity in African agriculture: insights from non-parametric and parametric analyses. J Clean Prod. 2025;513:145549. doi:10.1016/j.jclepro.2025.145549
3. BPS-Statistics Indonesia. Statistical Yearbook of Indonesia 2025. Jakarta: BPS-Statistics Indonesia; 2025.
Catalogue No.: 1101001.
Publication No.: 03200.25004.
ISSN/ISBN: 0126-2912.
https://www.bps.go.id/en/publication/2025/02/28/8cfe1a589ad3693396d3db9f/statistical-yearbook-of-indonesia-2025.html
4. Rashid M, Bari BS, Yusup Y, Kamaruddin MA, Khan N. A comprehensive review of crop yield prediction using machine learning approaches with special emphasis on palm oil yield prediction. IEEE Access. 2021;9:63406-63439. https://doi.org/10.1109/ACCESS.2021.3075159
5. Shafi U, Mumtaz R, Anwar Z, Ajmal MM, Khan MA, Mahmood Z, Qamar M, Jhanzab HM. Tackling food insecurity using remote sensing and machine learning-based crop yield prediction. IEEE Access. 2023;11:108640-108657. https://doi.org/10.1109/ACCESS.2023.3321020
6. Lu J, Li J, Fu H, Tang X, Liu Z, Chen H, Sun Y, Ning X. Deep learning for multi-source data-driven crop yield prediction in Northeast China. Agriculture. 2024;14(6):794. https://doi.org/10.3390/agriculture14060794
7. Demirhan H. Information Processing in Agriculture A deep learning framework for prediction of crop yield in Australia under the impact of climate change. Inf Process Agric. 2025;12(1):125–38. https://doi.org/10.1016/j.inpa.2024.04.004
8. Wolfert S, Ge L, Verdouw C, Bogaardt MJ. Big data in smart farming: a review. Agric Syst. 2017;153:69-80. https://doi.org/10.1016/j.agsy.2017.01.023
9. Gupta G, Kumar S. Applications of AI in precision agriculture. Discov Agric. 2025. https://doi.org/10.1007/s44279-025-00220-9
10. Jabed MA, Murad MAA. Crop yield prediction in agriculture: A comprehensive review of machine learning and deep learning approaches, with insights for future research and sustainability. Heliyon. 2024;10(24):e40836. https://doi.org/10.1016/j.heliyon.2024.e40836
11. Ismail UM, Bani-melhem K, Faizan M, Elnakar H. Results in Engineering Optimizing membrane bioreactor performance in wastewater treatment using machine learning and meta-heuristic techniques. Results Eng. 2025;25:104626. https://doi.org/10.1016/j.rineng.2025.104626
12. Uddien R, Alipour M, Rowell E, Balaji B, Watts A, Taciroglu E. FUELVISION : A multimodal data fusion and multimodel ensemble algorithm for wildfire fuels mapping . Int J Appl Earth Obs Geoinf. 2025;138:104436. https://doi.org/10.1016/j.jag.2025.104436
13. Yenkikar A, Prakash V, Bali M, Ara T. MethodsX An explainable AI-based hybrid machine learning model for interpretability and enhanced crop yield prediction. MethodsX. 2025;15:103442.

https://doi.org/10.1016/j.mex.2025.103442
14. Shafi U, Mumtaz R, Anwar Z, Ajmal MM, Khan MA, Mahmood Z.Tackling food insecurity using remote sensing and machine learning-based crop yield prediction. IEEE Access. 2023;11:108640-57. https://doi.org/10.1109/ACCESS.2023.3321020
15. Kulkarni VY, Sinha PK. Random Forest classifiers: a survey and future research directions. Int J Adv Comput. 2013;36(1):1144-53. https://adiwijaya.staff.telkomuniversity.ac.id/files/2014/02/Random-Forest-Classifiers_A-Survey-and-Future.pdf
16. Mohan RNVJ, Rayanoothala PS, Sree RP. Next-gen agriculture: integrating AI and XAI for precision crop yield predictions. Front Plant Sci. 2025;15:1451607. https://doi:10.3389/fpls.2024.1451607
17. Alsulamy S. Predicting construction delay risks in Saudi Arabian projects : A comparative analysis of CatBoost , XGBoost and LGBM. Expert Syst Appl. 2025;268:126268. https://doi.org/10.1016/j.eswa.2024.126268
18. Nguinabé J, Rockefeller R, Khan NM, Rodrigues PC. Bootstrap prediction intervals for the long short term memory (LSTM) recurrent neural network. Expert Syst Appl. 2025;284:127728. https://doi.org/10.1016/j.eswa.2025.127728
19. Ahmad H, Zafar SA, Naeem MK, Shokat S, Inam S, Rehman MAU, Naveed SA, Xu J, Khan MR. Impact of pre-anthesis drought stress on physiology, yield-related traits and drought-responsive genes in green super rice. Front Genet. 2022;13:832542. https://doi.org/10.3389/fgene.2022.832542
20. Rajbongshi A, Johora FT, Hossain A, Sarker MS, Rahman MH, Rahman MW, Alotaibi FT, Moni MA. Leveraging explainable AI for sustainable agriculture: a comprehensive review of recent advances. Artif Intell Rev. 2026;59(3):105. https://doi.org/10.1007/s10462-025-11459-5
21. Shastri S, Kumar S, Mansotra V, Salgotra R. Advancing crop recommendation system with supervised machine learning and explainable artificial intelligence. Scientific Reports. 2025;15:25498. https://doi.org/10.1038/s41598-025-07003-8
22. Xiong W, Xiao L, Han D, Yue W. A prediction model for water absorption in sublayers based on stacking ensemble learning method. Geoenergy Sci Eng. 2024;239:212896. https://doi.org/10.1016/j.geoen.2024.212896
23. Padhi J, Mishra K, Kumar A, Kumari S, Sethy PK, Nanthaamornphong A. Smart Agricultural Technology Enhancing paddy leaf disease diagnosis -a hybrid CNN model using simulated thermal imaging. Smart Agric Technol. 2025;10:100814. https://doi.org/10.1016/j.atech.2025.100814
24. Shawon SM, Ema FB, Mahi AK, Niha FL, Zubair HT. Crop yield prediction using machine learning: an extensive and systematic literature review. Smart Agric Technol. 2025;10:100718. https://doi.org/10.1016/j.atech.2024.100718
25. Senapaty MK, Ray A, Padhy N. IoT-Enabled Soil Nutrient Analysis and Crop Recommendation Model for Precision Agriculture. Computers. 2023;12(3):61. https://doi.org/10.3390/computers12030061
26. Akif M, Gunaydin S, Kavuncuoglu E, Cetin N, Sacilik K, Auat F. Model comparison and hyperparameter optimization for visible and near-infrared (Vis-NIR) spectral classification of dehydrated banana slices. Expert Syst Appl. 2025;283:127858. https://doi.org/10.1016/j.eswa.2025.127858
27. Hasan M, Marjan MA, Uddin MP, Afjal MI, Kardy S, Ma S, Nam Y. Ensemble machine learning-based recommendation system for effective prediction of suitable agricultural crop cultivation. Frontiers in Plant Science. 2023;14:1234555. https://doi.org/10.3389/fpls.2023.1234555
28. Islam MD, Di L, Qamer FM, Shrestha S, Guo L, Lin L, Mayer TJ, Phalke AR. Rapid Rice Yield Estimation Using Integrated Remote Sensing and Meteorological Data and Machine Learning. Remote Sensing. 2023;15(9):2374. https://doi.org/10.3390/rs15092374
29. Wang F, Hou X. Energy Geoscience Machine learning-based prediction of physical parameters in heterogeneous carbonate reservoirs using well log data. Energy Geosci. 2025;6(2):100383. https://doi.org/10.1016/j.engeos.2025.100383
30. Bhuiyan MD Shahriar Mahmud, Rafi MD AL, Rodrigues Gourab Nicholas, Mir MD Nazmul Hossain, Ishraq A, Mridha MF, Shin J.Deep learning for algorithmic trading : A systematic review of predictive models and optimization strategies. Array. 2025;26:100390. https://doi.org/10.1016/j.array.2025.100390
31. Gc S, Khan A, Horvath D, Sun X. Smart Agricultural Technology Evaluation of multispectral imaging for freeze damage assessment in strawberries using AI-based computer vision technology. Smart Agric Technol. 2025;10:100851. https://doi.org/10.1016/j.atech.2025.100851
32. Harsányi E. Predicting agricultural drought in Central Europe by using machine learning algorithms. J Agric Food Res. 2025;20:101783. https://doi.org/10.1016/j.jafr.2025.101783
33. Hamad K, Alotaibi E, Zeiada W, Al-Khateeb G, Abu Dabous S, Omar M, Mantha BRK, Arab MG, Merabtene T. Explainable artificial intelligence visions on incident duration using eXtreme Gradient Boosting and SHapley Additive exPlanations. Multimodal Transp . 2025;4(2):100209. https://doi.org/10.1016/j.multra.2025.100209
34. Chen HY, Sharma K, Sharma C, Sharma S. Integrating explainable artificial intelligence and blockchain to smart agriculture: Research prospects for decision making and improved security. Smart Agric Technol. 2023;6:100350. https://doi.org/10.1016/j.atech.2023.100350
35. Avazpour F, Hadian MR, Talebi A, Torabi Haghighi A. Impact of climate change on river flow, using a hybrid model of LARS_WG and LSTM: A case study in the Kashkan Basin. Results Eng. 2025;26:104956. https://doi.org/10.1016/j.rineng.2025.104956
36. Rashid MRA, Korim Md. AL E, Hasan M, Ali MS, Islam MM, Jabid T, Islam RU, Islam M. An Ensemble Learning Framework with Explainable AI for interpretable leaf disease detection. Array. 2025;26:100386. https://doi.org/10.1016/j.array.2025.100386
37. Sah S, Haldar D, Singh RN, Das B, Nain AS. Rice yield prediction through integration of biophysical parameters with SAR and optical remote sensing

data using machine learning models. Scientific Reports. 2024;14:22267. https://doi.org/10.1038/s41598-024-72624-4
38. Li W, Liu C, Xu Y, Niu C, Li R, Li M, Hu C, Tian L. An interpretable hybrid deep learning model for flood forecasting based on Transformer and LSTM. Journal of Hydrology: Regional Studies. 2024;54:101873. https://doi.org/10.1016/j.ejrh.2024.101873

How to Cite: Kusumodestoni RH, Arifin F, Nugraheni M. Optimized Weighted Fusion for Rice Productivity Prediction Using Multi-source Agricultural Data. Int Res J Multidiscip Scope. 2026;7(3):1530-1546.
DOI: 10.47857/irjms.2026.v07i03.012859