Vol. 8, No. 4, August 2026

ISI Journal of Information Systems and Informatics

Published By Asosiasi Doktor Sistem Informasi Indonesia

An Integrated XGBoost-SHAP Framework for XAI-based Analysis of Agroclimatic, Agronomic, and Operational Labor Proxy Affecting Block-Level Oil Palm Productivity

Hafis Ramadhan[1], Ahmad Jazuli[2], Esti Wijayanti[3]

[1, 2, 3] Informatics Engineering, Faculty of Engineering, Muria Kudus University, Kudus, Indonesia

Received:
February 19, 2026
Revised:
July 15, 2026
Accepted:
August 31, 2026
Published:
September 3, 2026

Corresponding Author:
Author Name*:
Hafis Ramadhan
Email*:
hafisramadhan111@gmail.com

DOI:
10.63158/journalisi.v8i4.1770

Abstract. Oil palm block-level productivity (Yield Per Hectare/YPH) relies on complex agronomic, climatic, and operational interactions, yet existing predictive models often neglect internal estate conditions. We developed an XGBoost-SHAP framework using 1,351 block-year observations (2018–2024) from a single plantation in Central Kalimantan, Indonesia. Features include plant age, land area, rainfall, and a labor ratio functioning as a land-based workload proxy. The baseline random-split testing achieved R^2 = 0.6276, MAE = 2.2747 Tons/Ha, RMSE = 3.3303 Tons/Ha, and MAPE = 28.73%. However, evaluating the model under grouped block and time-based validation schemes revealed limited spatial and temporal transferability. SHAP analysis identified plant age (47.51%), land area (14.32%), annual rainfall (10.87%), and labor ratio (8.51%) as primary yield drivers, revealing non-linear interaction patterns. Rather than a fully autonomous tool, this framework provides a transparent interpretive heuristic and decision-support prototype to evaluate block-level operational strategies. By quantifying internal operational variables alongside agronomic factors, this study contributes valuable insights for precision agriculture, suggesting future work validate dashboard heuristics with estate managers and incorporate granular fertilization and harvesting logs to further enhance predictive robustness.

Keywords: Block-Level Prediction, CRISP-DM, Explainable AI, Oil Palm Yield Prediction, Precision Agriculture, SHAP, XGBoost

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

1. INTRODUCTION

The oil palm plantation industry plays a strategic role in the Indonesian economy, with ongoing management interventions aimed at optimizing field productivity and agronomic sustainability [1]. However, realized Yield Per Hectare (YPH) exhibits dynamic variability and is nonlinearly influenced by climate anomalies [2], while also showing substantial variation across plantation environments [3]. Within the agroclimatic and agronomic domains, extreme weather events and biological aging significantly impair tree productivity [4], [5]. Furthermore, empirical studies confirm that substantial yield gaps in oil palm plantations are driven by agronomic management deficiencies, improper fertilizer placement, and sub-optimal fertilization regimes that prevent crops from reaching their potential yield capacity [6], [7], [8]. In the operational domain, labor allocation and workforce availability are closely associated with harvesting productivity [9], [10], while excessive physical workload also significantly affects harvester operational performance [11]. Moreover, extended harvesting intervals and delayed rotation rounds directly increase in-field crop losses and reduce bunch recovery [12], [13].

To anticipate these complex interactions of field factors, modern, measurable precision agriculture systems utilizing Machine Learning (ML) algorithms have been increasingly adopted [14]. To ensure optimal predictive capability, continuous refinements in data preparation and feature representation remain a critical foundation for minimizing noise in multi-attribute agricultural data configurations.

So far, the application of tree-based Machine Learning integrated with Explainable Artificial Intelligence (XAI) for oil palm yield prediction has been demonstrated by Elwirehardja et al. [15] who employed XGBoost and SHAP to analyze agrometeorological variables together with plantation characteristics and historical yield records, achieving a high predictive performance (R^2 = 0.855, RMSE = 1.911). However, several limitations remain. First, the model did not explicitly incorporate internal estate operational variables, such as harvester workload or weather-related operational disruption. Second, the authors acknowledged that rainfall and rainy-day data were recorded at the regional level, resulting in identical meteorological values across multiple plantation fields and limiting the spatial resolution of environmental information. Consequently, although the study successfully identified influential variables through SHAP, it did not explicitly integrate internal estate operational dynamics into an operational modeling framework for block-level managerial decision support.

Other studies applying approaches with more limited variables reported a more moderate range of accuracy; weather and soil moisture-based approaches recorded an R^2 = 0.60 – 0.63 [16], while methods utilizing vegetation and moisture indices based on Landsat satellite imagery successfully demonstrated the feasibility of using remote sensing data to predict harvest productivity spatially [17]. Although multisource satellite imagery approaches, such as Landsat and MODIS, can generate yield estimates at the plantation block scale with varying degrees of accuracy [18], many existing predictive models still focus on macro external factors while neglecting internal operational conditions of the estate, such as harvester workload and field weather disruptions [14].

Motivated by these limitations, this study contributes to the literature by developing a multivariate XGBoost-SHAP modeling and interpretation framework that synthesizes agroclimatic, agronomic, and operational factors. In particular, a land-based harvester workload proxy (Labor Ratio) and an Operational Disruption Index are introduced to represent internal estate operational conditions that were not explicitly represented in prior frameworks [15]. By integrating these operational variables alongside environmental and biological drivers, the proposed framework not only predicts block-level oil palm productivity but also provides transparent, actionable decision support for estate managers, assistant managers, resident agronomists, and harvest supervisors. Importantly, the empirical scope, feature interactions, and operational heuristics derived in this study are specifically contextualized to the observed plantation estate in Central Kalimantan and are not intended as broad generalizations for all Indonesian oil palm plantations with differing soil types, topographies, or microclimatic conditions. Specifically, the primary objectives of this research are to: (1) evaluate the predictive performance of a 9-feature multivariate model using the XGBoost Regressor; (2) analyze the direction, magnitude, and non-linear interactions of these features using SHAP interpretation; and (3) formulate actionable, XAI-based managerial insights to support operational decision-making at the plantation block level.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

2. METHODS

This study applies the CRISP-DM (Cross Industry Standard Process for Data Mining) framework [19] to build a multivariate XGBoost-SHAP modeling and interpretation framework, which consists of six phases: business understanding, data understanding, data preparation, modeling, evaluation, and implementation. This integrated framework was selected for its capability to seamlessly combine XGBoost predictive modeling with SHAP feature interpretation, ensuring that plantation logic is embedded into every phase of the research. Consequently, each data processing step and model output explanation directly reflects actual operational conditions in the field.

The research initiates with data collection and understanding sourced from internal company records, followed by data preparation through comprehensive feature engineering procedures. Once the dataset is finalized, the study proceeds to the modeling stage using the XGBoost algorithm, which is subsequently assessed through model performance evaluation. The final phase (implementation) is executed by leveraging model interpretation via the SHAP (Shapley Additive exPlanations) method to derive actionable insights, with the entire operational workflow systematically summarized in Figure 1.

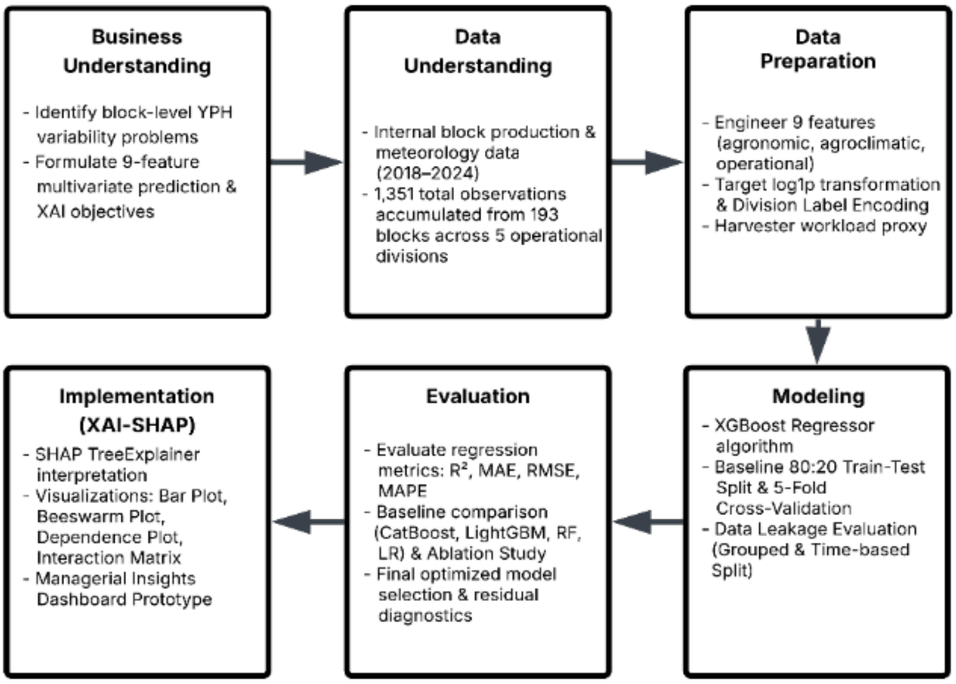

Figure 1. Research Flowchart Based on the CRISP-DM Life Cycle.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

2.1 Business Understanding

The business understanding phase aims to identify and systematically define operational problems within oil palm plantations, focusing specifically on quantifying block-level YPH variability problems. In this context, the evaluation addresses the high variability of Yield Per Hectare (YPH) at the block level—a fluctuation that remains difficult to explain comprehensively using a single variable alone [3]. Driven by this problem, this research is designed to achieve three primary objectives: (1) evaluate the predictive performance of a 9-feature multivariate model using the XGBoost Regressor, (2) analyze the direction, magnitude, and non-linear interactions among features using SHAP interpretation, and (3) formulate XAI-based managerial recommendations to support operational decision-making at the plantation block level.

2.2 Data Understanding

The data utilized in this study were sourced from the internal archives of an oil palm plantation company located in Central Kalimantan, Indonesia, spanning a historical observation period from 2018 to 2024. This dataset comprises two primary data groups: block-level production data—which includes YPH values (measured in Tons/Ha), land area, planting year, and operational estate divisions—and meteorological data consisting of monthly rainfall and monthly rainy days.

The data integration process was executed by merging these two data streams using a harvest year key as the relational link, resulting in a finalized dataset containing 1,351 block-year observations. This structured data sample encompasses 193 blocks accumulated across five operational divisions, namely Divisions 1, 2, 4, 5, and 6. Division 3 was intentionally excluded from the data pipeline because it represents a newly developed operational area that lacks a complete historical production track record for the designated 2018–2024 observation period. This exclusion does not introduce systemic sampling bias; rather, it ensures dataset integrity by avoiding severe panel imbalance and structural zero-value truncation, while preserving the representation of mature, producing block yield dynamics across the remaining divisions. Furthermore, observations with a YPH value of zero in the included divisions were intentionally retained within the dataset to accurately reflect actual operational realities in the historical data.

ISI Journal of Information Systems and Informatics

Published By Asosiasi Doktor Sistem Informasi Indonesia

2.3 Data Preparation

The data preparation phase involved converting the plantation production data from a wide format into a long format, ensuring that each row represents a single block-year observation. A feature engineering procedure was executed to construct nine predictor variables spanning three primary plantation domains: agronomic, agroclimatic, and operational [2], [3]. A comprehensive summary of all nine predictor features, including their domain categories, measurement units, calculation formulas, and data sources, is presented in Table 1.

Table 1. Summary of Predictor Features across Agronomic, Agroclimatic, and Operational Domains

Domain Category	Variable Name	Unit	Calculation Method / Definition	Data Source
Agronomic	Plant Age	Years	Observation Year – Planting Year	Internal Records
Agronomic	Land Area	Hectares (Ha)	Physical block land area measurement	Internal Records
Agroclimatic	Annual Rainfall	mm	Cumulative sum of monthly rainfall in a year	Rainfall Log
Agroclimatic	Annual Rainy Days	Days	Cumulative count of rainy days (≥ mm/day) in a year	Rainfall Log
Agroclimatic	Dry Months	Months	Count of months in a year with monthly rainfall < 60 mm	Derived Feature
Agroclimatic	Extreme Rain Months	Months	Count of months in a year with > 22 rainy days	Derived Feature
Operational	Operational Disruption Index (ODI)	Index Score	Sum of Extreme Rain Months and Dry Months	Derived Feature
Operational	Labor Ratio	Ha/Worker	Land Area ÷ Estimated Harvesters (assuming a	Derived Feature

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

Domain Category	Variable Name	Unit	Calculation Method / Definition	Data Source
			standard operational workload capacity of 8 Ha per harvester)	
Management	Estate Division	Categorical ID	Integer encoded division identifier (Divisions 1, 2, 4, 5, 6)	Internal Records

Regarding feature representation, Plant Age was utilized as a continuous numerical variable (in years) during model training to enable XGBoost to learn granular, non-linear continuous threshold effects. The discrete categorical grouping into Young (3–7 years), Prime (8–18 years), and Old (>18 years) was exclusively employed during initial Exploratory Data Analysis (EDA) for descriptive aggregation. Based on the distribution of the historical data utilized in this study, the 2018–2024 records indicated an absence of plantation blocks falling into the Old category; thus, all historical observation samples realistically occupied only the Young and Prime phases.

The Labor Ratio—or the hectare-to-harvester ratio—was constructed using a land area-based proxy approach, as explicit labor distribution information at the block level was unavailable in the research dataset. The proxy for the number of harvesters was computationally estimated based on an operational capacity assumption of 8 hectares per harvester. Consequently, the ratio was calculated as the proportion of the block land area relative to the estimated number of harvesters to adequately represent the operational workload at the block level. Modeling labor operational capacity via this land-based proxy allows the framework to systematically bridge internal operational workloads with external agroclimatic factors for block-level yield prediction.

The Operational Disruption Index (ODI) was developed to represent potential disruptions to harvesters' operational activities caused by adverse weather conditions. The index value was calculated as the sum of extreme rainy months and dry months using the formula as shown in Equation 1.

$$ODI = \text{Extreme Rain Months} + \text{Dry Months} \tag{1}$$

These zero-YPH observations do not represent missing values or immature blocks; rather, they represent real non-harvest events caused by severe localized operational disruptions (e.g., extreme waterlogging and temporary harvesting access blockages).

The categorical variable estate division was transformed into numerical values using the Label Encoding method. To address potential concerns regarding artificial split thresholds, we methodologically compared Label Encoding versus One-Hot Encoding (OHE). Empirical evaluation demonstrated that OHE yielded a testing R^2 of 0.6430 and an MAE of 2.1557 Tons/Ha, providing only a marginal performance difference (ΔR^2 = +0.0154) compared to Label Encoding (R^2 = 0.6276, MAE = 2.2747 Tons/Ha). Additionally, a 5-fold cross-validation robustness check confirmed similar stability (OHE CV R^2 = 0.5835 vs Label CV R^2 = 0.5721). Because tree-based ensemble algorithms like XGBoost perform recursive binary partition splits on feature values without assuming continuous spatial ordering or distance metrics between integer labels, they efficiently handle label-encoded features. Consequently, Label Encoding was retained to prevent sparse matrix expansion (increasing the input array from 9 to 12 features) across our modest tabular dataset (n = 1,351), thereby preserving overall SHAP feature attribution integrity. Finally, all preprocessing steps, including target scaling via the $\log 1\mathrm{p}$ function and data encoding, were strictly fitted only on the training set within each respective validation scheme to prevent data leakage. The model's predicted values were transformed back to their original scale using the expm1 inverse function immediately before the evaluation metrics were computed.

2.4 Modeling

Predictive modeling in this study utilizes the Extreme Gradient Boosting (XGBoost) Regressor algorithm. This algorithm minimizes a regularized objective function combining convex loss and tree complexity penalties, formulated as shown in Equation 2 [20]:

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \tag{2}$$

Where l represents the differentiable convex loss function measuring prediction error between the actual yield (y_i) and the predicted yield at step $t-1\left(\hat{y}_i{}^{(t-1)}\right)$, $f_t(x_i)$ is the functional mapping for the new t-th tree, n is the number of observations, and $\Omega(f_t)$

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

penalizes model complexity to prevent overfitting across tabular plantation attributes [20]. The complexity penalty term is defined as shown in Equation 3.

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \tag{3}$$

Where T denotes the number of terminal leaf nodes, w_j represents the leaf weight, γ is the complexity penalty coefficient, and λ is the L2 regularization factor on leaf weights.

The model training phase was executed by partitioning the dataset into an 80% training set and a 20% testing set—an 80:20 train–test split—which yielded 1,080 training samples and 271 testing samples, using a fixed random_state of 42. To rigorously evaluate potential data leakage risks inherent in repeated block-year panel observations, two additional validation schemes were implemented alongside the baseline random split: (1) a Grouped Split by Block, where all historical observations of any given block were strictly isolated into either the training or testing set ($n_{test} = 273$) to prevent spatial overlap; and (2) a Time-based Split, partitioning data into a 2018–2022 training period ($n_{train} = 965$) and a 2023–2024 testing period ($n_{test} = 386$). The time-based scheme simulates a realistic operational scenario for forecasting future block-level yield without temporal information leakage. Subsequently, the stability and generalization capability of the model were rigorously evaluated using a 5-fold cross-validation scheme, considering potential spatial and temporal dependence structures within the block-year panel dataset [21].

The final XGBoost hyperparameter configuration (n_estimators = 2500, learning_rate = 0.04, max_depth = 4, subsample = 0.85, colsample_bytree = 0.85, reg_alpha = 0.5, and reg_lambda = 1.5) was established through manual iterative experimental tuning combined with a bounded grid search. The search space included ranges for learning_rate (0.01–0.1), max_depth (3–8), n_estimators (500–3000), subsample and colsample_bytree (0.6–1.0), and regularization parameters (0.0–2.0). Approximately 50 iterative trials were conducted, evaluated via 5-fold cross-validation on the training set using RMSE reduction as the primary selection criterion to achieve an optimal balance between prediction accuracy and model regularization against overfitting. To prevent data leakage during early stopping optimization, an internal 15% validation partition was strictly separated from the training data within the cross-validation and comparative split loops.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

Throughout the training process, an early stopping mechanism with a patience threshold of 100 iterations was applied to halt optimization if no further performance improvement was observed on the validation set. All experimental workflows and model evaluations were executed on a Linux x86_64 computing platform using Python 3.12.13, scikit-learn 1.6.1, XGBoost 3.2.0, SHAP 0.52.0, LightGBM 4.6.0, and CatBoost 1.2.10.

2.5 Evaluation

The evaluation phase measures the continuous estimation error of the predictive model based on experimental comparisons derived from object-oriented programming functions. The regression evaluation metrics are computed on the original scale of Tons/Ha after applying the expm1 inverse transformation function, using four primary evaluation metrics:

1) Coefficient of Determination (R^2)

Used to measure the proportion of variance in the target productivity variable that can be explained by the predictive model. This value is calculated using Equation 4.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{4}$$

2) Mean Absolute Error (MAE)

MAE used to calculate the average absolute difference between the actual productivity values and the model predictions, formulated as shown in Equation 5.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{5}$$

3) Root Mean Square Error (RMSE)

RMSE used to measure the magnitude of the prediction error by assigning a higher penalty to larger errors through the square root of the average squared differences, as shown in Equation 6.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{6}$$

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

4) Mean Absolute Percentage Error (MAPE):

MAPE used to evaluate the relative percentage error between actual productivity and model predictions, providing a scale-independent metric for cross-baseline comparison, as shown in Equation 7.

$$\mathrm{MAPE} = \frac{100\%}{n_{\text{valid}}} \sum_{i \in S_{\text{valid}}} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{7}$$

In these evaluation metric formulations, y_i represents the actual productivity value, $\hat{y}_i$ indicates the predicted productivity value generated by the model, $\bar{y}$ signifies the mean productivity value of the testing dataset, and n represents the total number of observations in the testing set ($n = 271$). For the MAPE calculation, it is mathematically undefined when the actual value is zero ($y_i = 0$). Therefore, observations where actual $y_i = 0$ (representing real non-harvest events) were explicitly excluded from the valid summation subset $S_{valid}(n_{valid} \leq n)$ solely during the MAPE computation to prevent division-by-zero errors. These four metrics are utilized complementarily to evaluate the model performance and serve as the baseline for selecting the optimal model for the implementation phase. To evaluate the predictive superiority of the proposed framework, performance metrics were benchmarked against several baseline algorithms (CatBoost, LightGBM, Random Forest, Linear Regression, and a Simple Baseline). These comparative baselines were executed using their default hyperparameter configurations without additional tuning to serve as standard out-of-the-box benchmarks against the optimally tuned XGBoost model.

2.6 Implementation XAI (SHAP)

The implementation phase applies Shapley Additive exPlanations (SHAP) to interpret feature contributions from the optimized model object. SHAP computes feature attributions as Shapley values (ϕ_i) based on additive feature attribution methods, formulated as shown in Equation 8 [22]:

$$g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i \, z_i' \tag{8}$$

Where g represents the local explanation model, $z' \in \{0,1\}^M$ is a binary coalition vector of present features, M and is the maximum number of input features, ϕ_0 is the base expected value of the model output, and $\phi_i \in \mathbb{R}$ quantifies the marginal yield contribution of feature i across block-level predictions [22], [23]. This study integrates the SHAP TreeExplainer approach, which is specifically designed to efficiently interpret

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

decision tree-based model structures by calculating Shapley values as a metric of each feature's contribution to the prediction outputs across the 271 testing data samples.

The feature importance bar plot illustrates the global contribution level of each feature toward the model predictions, while the beeswarm plot analyzes the direction and magnitude of each feature impact on the predicted values. Furthermore, feature attributions are computed following the unified Shapley framework [22], with foundational game-theoretic principles outlined in [23]. Dependence plots are utilized to evaluate non-linear response patterns and feature relationships at the block level, supported by recent applications of XAI in agricultural crop yield prediction [24]. The resulting interpretation outputs are then utilized as the foundational baseline for formulating managerial insights to support operational decision-making and strategy refinement within the plantation.

3. RESULTS AND DISCUSSION

This section presents the empirical results and analytical interpretations derived from applying the multivariate XGBoost-SHAP framework to the 2018–2024 oil palm plantation dataset. The discussion initiates with an exploratory data analysis of historical production and agroclimatic trends, followed by a comprehensive performance evaluation of the predictive model and baseline comparisons. Next, a SHAP-based feature importance interpretation uncovers variable attributions, concluding with residual diagnostics, data leakage discussions, and an XAI-based managerial insights prototype.

3.1. Exploratory Data Analysis

Plantation productivity over the 2018 to 2024 period exhibited dynamic fluctuations, with the average Yield Per Hectare (YPH) increasing progressively from 3.71 Tons/Ha to a peak of 9.08 Tons/Ha in 2021. As illustrated in Figure 2, this upward trend underwent a temporary correction, dropping to 6.22 Tons/Ha in 2022, before ultimately surging to its highest recorded value of 12.10 Tons/Ha in 2024.

The yield reduction observed in 2022 indicates the presence of a biological lag mechanism triggered by weather anomalies. The highly extreme accumulation of rainfall during the preceding period—reaching a record high of 6,299 mm in 2020—can induce

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

excessive waterlogging and physiological stress in the root systems. The cumulative impact of such conditions typically manifests in subsequent harvest seasons, aligning with findings that oil palm productivity is highly responsive to climatic anomalies occurring 1 to 2 years prior to harvest due to physiological lag mechanisms in floral initiation and fruit bunch development [25]. This phenomenon is clearly evident in Division 4, where several previously productive blocks recorded a YPH of zero in 2022 and only began to recover in 2023.

The distribution of yield based on plant age phases demonstrates that the prime phase group (8–18 years) consistently generated the highest median YPH across all operational estate divisions. The productivity range of prime phase blocks outperformed that of young phase blocks (3–7 years), whose production capacity remains physiologically limited. This pattern is consistent with the existence of a peak production phase in oil palm crops as widely reported in prior agronomic literature [4].

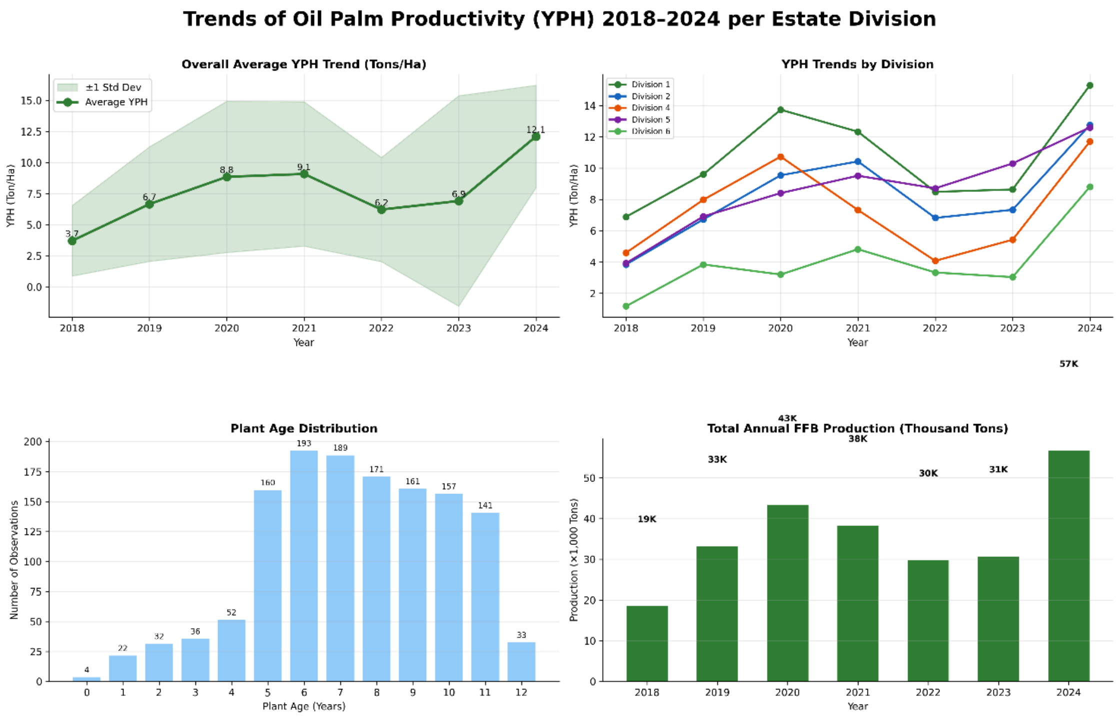

Figure 2. Annual YPH trends and FFB production (2018–2024), YPH trends by division, and plant age distribution.

A comprehensive agroclimatic analysis is presented in Figure 3. Through the visualization of annual rainfall from 2018 to 2024, it is detected that the highest accumulation of

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

rainfall occurred in 2020, reaching a total of 6,299 mm. This condition indicates that highly intense and extreme rainfall events took place in 2020. Meanwhile, in terms of the frequency of rainy days, the peak conversely occurred in 2024. This confirms that 2024 experienced the most frequent rainfall events—occurring almost daily—even though its cumulative total volume did not surpass the record set in 2020. Furthermore, the monthly heatmap visualization reveals a consistent pattern of climate anomalies, where rainfall intensity and the number of rainy days sharply increase every December across all observed years. This recurring year-end extreme pattern serves as a significant signal regarding the high risk of harvesting operational disruptions during that specific period.

Based on the scatter plot correlating YPH with annual rainfall in Figure 3, an increasing trend in productivity is observed to align with higher rainfall volumes. Within the observation range of this plantation ecosystem (4,000–6,300 mm), elevated annual rainfall is consistently associated with higher average YPH values. This empirical finding is in line with previous research reporting a strong positive correlation between rainfall and FFB production in the Central Kalimantan region [26]. These results indicate that adequate water availability constitutes a crucial factor in supporting the formation and development of oil palm fresh fruit bunches.

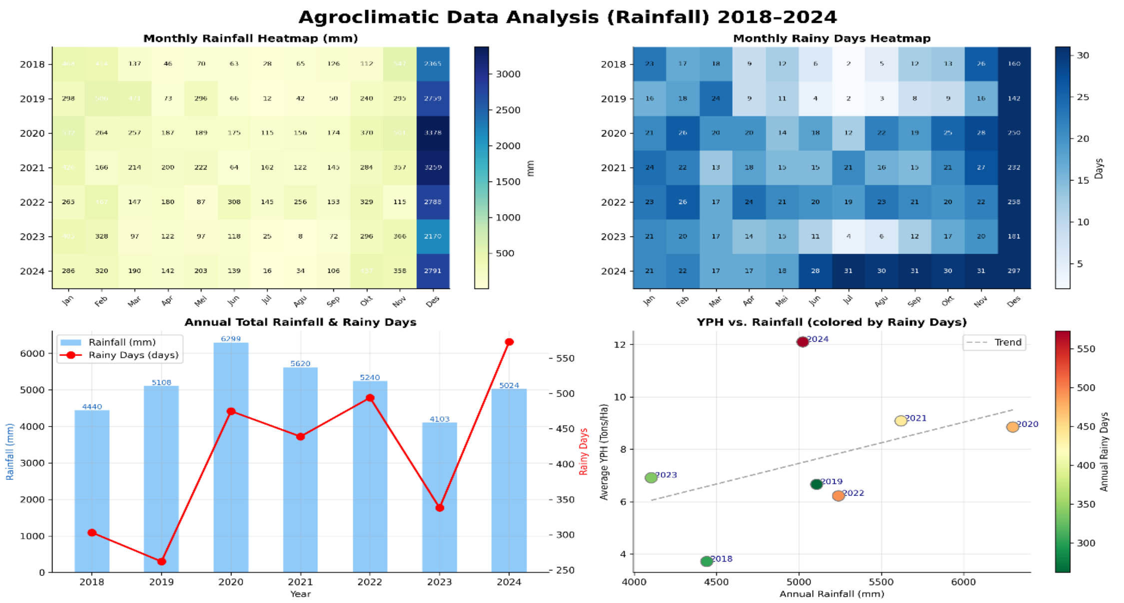

Figure 3. Monthly rainfall and rainy days heatmaps, annual trends of total rainfall and rainy days (2018–2024), and scatter plot of YPH vs. rainfall.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

The Pearson correlation matrix (Figure 4) evaluates the linear relationships of 11 candidate predictor features during the initial stage of exploratory data analysis, including additional agroclimatic variables such as Max Monthly Rainfall and Wet Months, before they were ultimately filtered down into the 9 final operational features. The analytical results confirm that plant age exhibits the largest positive correlation value toward the target YPH ($r = 0.530$). Although there is an indication of a linear relationship for this specific variable, overall, none of the eleven single features are capable of absolutely dominating the target linearly ($r > 0.8$). This condition underscores the necessity of employing a non-linear ensemble-based machine learning algorithm to capture the complex interactions among variables that cannot be adequately explained by standard linear correlation analysis.

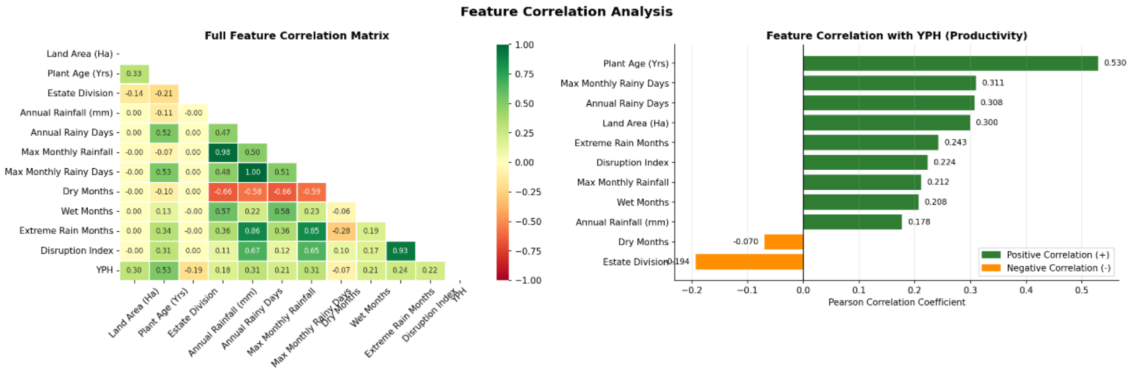

Figure 4. Pearson Correlation Matrix of All Features and YPH

3.2. Evaluation of XGBoost Model Performance

Upon evaluating the finalized XGBoost model on the baseline 80:20 partition, the training set ($n = 1{,}080$) achieved an R^2 of 0.7555, MAE of 1.5726 Tons/Ha, RMSE of 2.9959 Tons/Ha, and MAPE of 20.74%. On the hold-out testing set ($n = 271$), the model achieved an R^2 of 0.6276, MAE of 2.2747 Tons/Ha, RMSE of 3.3303 Tons/Ha, and MAPE of 28.73% (Table 2 and Table 3). The moderate gap between training and testing performance confirms that the applied regularization parameters successfully prevented severe overfitting across tabular block-year attributes. Furthermore, to evaluate model stability across fold partitions, 5-fold cross-validation on the training set yielded a mean R^2 of 0.5721 ± 0.1844, MAE of 2.2067 ± 0.2087 Tons/Ha, RMSE of 3.8478 ± 1.5334 Tons/Ha, and MAPE of $76.78\% \pm 20.34\%$. The testing RMSE value (3.3303 Tons/Ha), which is slightly higher than the MAE (2.2747 Tons/Ha), indicates localized variance and larger prediction errors

ISI Journal of Information Systems and Informatics

Published By Asosiasi Doktor Sistem Informasi Indonesia

in specific operational blocks, largely influenced by retaining real operational zero-YPH observations.

However, a comparative evaluation across the three validation schemes reveals critical constraints regarding model generalization. While the random split configuration effectively explains cross-sectional operational variance within concurrent years, isolating entire blocks in the Grouped Split led to a noticeable drop in predictive power (R^2 = 0.2300, MAE = 3.1716 Tons/Ha), indicating localized spatial dependency. Furthermore, evaluating under a Time-based Split resulted in an R^2 of -0.1106. This performance decline underscores severe inter-annual climate shifts and biological lag effects following the 2020–2022 weather anomalies, confirming that the baseline random split carries spatial and temporal leakage risks that limit cross-temporal generalization without continuous multi-year recalibration.

Table 2. Comparative Performance across Validation Schemes for Data Leakage Evaluation

Validation Scheme	**Data Partitioning Strategy**	R^2	**MAE (Tons/Ha)**	**RMSE (Tons/Ha)**	**Test Samples (n)**	**MAPE (%)**
Random Split (Baseline)	Random 80:20 per block-year observation	0.6276	2.2747	3.3303	271	28.73%
Grouped Split by Block	Strict block isolation (no spatial overlap)	0.2300	3.1716	7.2709	273	43.55%
Time-based Split	Train: 2018–2022 Test: 2023–2024	-0.1106	4.8479	7.5095	386	51.00%

A comparative evaluation across the three validation schemes (Table 2) reveals critical insights into the spatial and temporal dynamics of the dataset. While the baseline random split achieved an R^2 of 0.6276 with an MAE of 2.2747 Tons/Ha, isolating entire blocks in the Grouped Split led to a drop in predictive power (R^2 = 0.2300, MAE = 3.1716 Tons/Ha). This discrepancy confirms the presence of block-level spatial dependencies, indicating that the baseline model partially leverages unobserved block-specific identifiers (e.g., specific micro-soil conditions or localized management habits) to refine its predictions.

Furthermore, evaluating the model under a Time-based Split (training on 2018–2022 and testing on 2023–2024) resulted in an R^2 of -0.1106. This decline highlights a pronounced distributional shift in agroclimatic conditions between the two time periods—most notably the extreme rainfall anomalies in 2024 and the biological lag effects following the 2020–2022 climate disruptions. These findings demonstrate that while the random split configuration is highly effective for explaining cross-sectional operational variance within concurrent years, multi-year yield forecasting requires incorporating longer climate cycles and adaptive temporal features to account for inter-annual climate shifts. To evaluate the necessity and predictive performance of the proposed XGBoost architecture, metrics were benchmarked against several baseline regression algorithms using the same 80:20 train-test split (Table 3). As presented in Table 3, XGBoost achieved the highest variance explanation (R^2 = 0.6276) and lowest large-error penalty (RMSE = 3.3303 Tons/Ha), whereas CatBoost exhibited a marginally lower MAE (2.2630 Tons/Ha) and MAPE (28.58%). Because their test-set metrics and cross-validation standard deviations are closely aligned, a paired absolute error analysis was performed (n = 271). The mean absolute error difference between XGBoost and CatBoost was negligible (+0.0117 Tons/Ha). Both a Wilcoxon signed-rank test (W = 17,995.0, p = 0.7374) and a paired t-test (t = 0.2874, p = 0.7741, Cohen's d = 0.017) confirmed that the performance difference between XGBoost and CatBoost is not statistically significant. XGBoost was selected for the final interpretation pipeline due to its seamless integration with TreeSHAP and optimal capture of non-linear interaction thresholds.

Table 3. Performance Benchmark Comparison Against Baseline Algorithms (Testing Set Metrics $\pm$ 5-Fold CV Standard Deviation)

Algorithm / Model	R^2 (Score)	MAE (Tons/Ha)	RMSE (Tons/Ha)	MAPE (%)
XGBoost (Proposed, Tuned)	0.6276±0.1844	2.2747±0.2087	3.3303±1.5334	28.73%±1.48%
CatBoost	0.6235±0.1752	2.2630±0.1885	3.3489±1.5076	28.58%±1.18%
LightGBM	0.6020±0.1756	2.3311±0.1986	3.4430±1.4984	29.44%±1.44%
Random Forest	0.5648±0.1598	2.6133±0.1732	3.6004±1.4481	33.00%±0.92%

Algorithm / Model	R^2 (Score)	MAE (Tons/Ha)	RMSE (Tons/Ha)	MAPE (%)
Linear Regression	0.4177±0.1315	3.1274±0.2039	4.1646±1.3562	39.50%±1.85%
Simple Baseline (Plant Age + Rainfall)	0.2519±0.1177	3.5858±0.2404	4.7206±1.2714	45.29%±2.18%

To evaluate the specific predictive contribution of internal operational variables, an ablation study was conducted by retraining the model without the operational feature block (Land Area, Operational Disruption Index, and Labor Ratio). As shown in Table 4, removing the operational feature block caused the testing R^2 to drop from 0.6276 to 0.5480 and the training R^2 to decrease substantially from 0.7555 to 0.4763. Furthermore, error metrics expanded across both partitions; testing MAE increased from 2.2747 Tons/Ha to 2.6525 Tons/Ha (MAPE from 28.73% to 33.50%), while training MAE increased from 1.5726 Tons/Ha to 2.4155 Tons/Ha (MAPE from 20.74% to 31.90%). These empirical comparisons confirm that operational variables provide essential explanatory power for block-level yield prediction.

Table 4. Ablation Study Comparing Performance With and Without Operational Variables (Testing Set n = 271)

Model Configuration	Dataset	R^2	MAE (Tons/Ha)	RMSE (Tons/Ha)	MAPE (%)
All Model 9 Features (Proposed)	Train	0.7555	1.5726	2.9959	20.74%
	Test	0.6276	2.2747	3.3303	28.73%
Ablated Model 6 Features	Train	0.4763	2.4155	4.3846	31.90%
	Test	0.5480	2.6525	3.6691	33.50%

The determination values and error rates obtained are also influenced by modeling limitations that focus on internal operational data without additional external agronomic variables. Consequently, the testing R^2 value of 0.6276 is lower than in previous studies reporting performance of up to R^2 = 0.855 using comprehensive agrometeorological

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

variables [15], but it remains comparable to benchmarks utilizing limited environmental features (R^2 = 0.60-0.63) [16] as well as satellite-based yield estimation models [17]. Despite the high field data variability caused by extreme agroclimatic factors, this optimized XGBoost model framework still demonstrates a reasonably stable generalization accuracy.

As illustrated in Figure 5, the empirical relationship between actual and predicted Yield Per Hectare (YPH) across the test set (n = 271) demonstrates a balanced clustering along the ideal 1:1 reference line. The XGBoost model exhibits consistent predictive accuracy across low, medium, and high yield regimes (R^2 = 0.6276, MAE = 2.2747 Tons/Ha), with minimal systematic bias across operational block conditions.

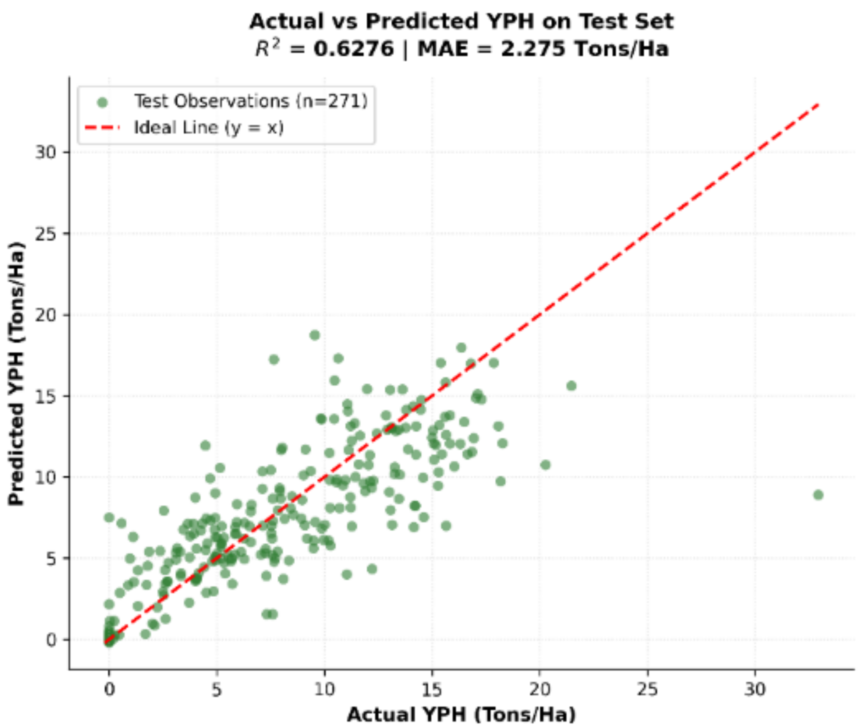

Figure 5. Scatter plot of actual versus predicted Yield Per Hectare (YPH)

3.3. Residual Diagnostics and Stratified Error Analysis

To evaluate model error behavior and detect potential systemic bias, residual diagnostic plots and stratified error analyses were conducted across operational dimensions (Figure 6). The residual scatter plot demonstrates a balanced homoscedastic distribution

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

symmetrically centered around zero across predicted Yield Per Hectare (YPH) values, confirming the absence of severe non-linear bias in model errors.

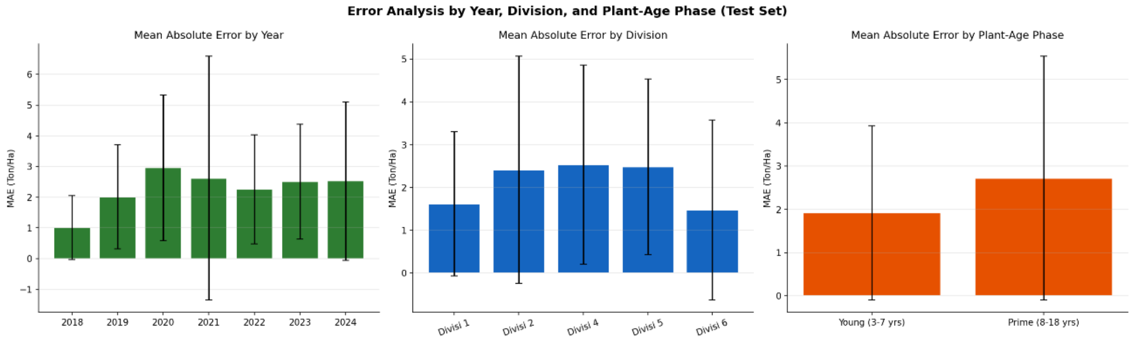

Figure 6. Stratified Mean Absolute Error (MAE) analysis with standard deviation error bars across Observation Years, Estate Divisions, and Plant-Age Growth Phases.

Stratified error analysis with standard deviation bounds across temporal, spatial, and agronomic phases reveals distinct operational error patterns (Figure 6):

1) Temporal Dimension (By Year): Prediction errors (MAE) were slightly higher in 2021 (MAE = 2.651 Tons/Ha) and 2022 (MAE = 2.584 Tons/Ha) compared to other observation years. This increased error variance corresponds to extreme rainfall anomalies and post-pandemic harvesting operational disruptions recorded during those specific periods.
2) Spatial Dimension (By Estate Division): Error variations across estate divisions remained stable overall, with Division 6 exhibiting the lowest MAE ($\approx$ 1.85 Tons/Ha) due to consistent block maturity, whereas Division 4 recorded slightly higher error variance due to localized flooding and temporary road access restrictions during wet months.
3) Agronomic Dimension (By Plant-Age Phase): Blocks in the prime yielding phase (TM-2 / TM-3) exhibited higher absolute yield variability and slightly larger MAE compared to immature/young blocks (TBM / TM-1). This reflects the higher absolute production baseline of mature oil palms, where minor operational harvesting variations result in larger absolute Ton/Ha deviations.

Furthermore, an analysis of zero-YPH records revealed that zero-value entries accounted for a minimal fraction of the dataset, representing genuine non-harvested events caused by extreme localized waterlogging or temporary access blockages rather than biological

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

crop failures. Removing zero-YPH rows during evaluation slightly adjusted the testing R^2 from 0.6276 to 0.6310, demonstrating that the core XGBoost framework remains resilient to operational outliers. To identify structural outlier patterns and contextualize prediction deviations, the top 10 block-year observations with the largest absolute prediction errors are detailed in Table 5.

Table 5. Top 10 Block-Year Observations with the Largest Absolute Prediction Error (Test Set)

Year	Block ID	Division	Phase	Actual YPH	Pred. YPH	Abs. Error
2021	BP50-13A1	Division 2	Prime	32.93	8.89	24.04
2024	BO36-13A2	Division 2	Prime	7.65	17.23	9.58
2020	BM29-13A2	Division 2	Young	20.27	10.74	9.53
2024	BQ43-13A1	Division 6	Prime	9.55	18.73	9.18
2024	BP32-13A1	Division 4	Prime	15.65	6.99	8.66
2020	BM25-13A2	Division 2	Young	18.18	9.72	8.46
2022	BP31-13B2	Division 4	Prime	12.21	4.33	7.88
2021	BN34-15C2	Division 5	Young	0.00	7.49	7.49
2020	BQ28-13A1	Division 5	Young	4.47	11.92	7.45
2021	BQ46-13A1	Division 6	Prime	14.15	6.91	7.24

As detailed in Table 5, the highest residual deviations were primarily concentrated in specific mature blocks during the 2021–2022 anomaly period—particularly within Division 4. These severe prediction errors directly correspond to localized extreme waterlogging and temporary harvesting access restrictions, where blocks recorded zero or near-zero YPH despite favorable agronomic crop age profiles.

3.4. SHAP Feature Importance and Non-linear Interpretation

All Shap values presented in this section were calculated on model predictions mapped back to the original target scale (Tons/Ha). Consequently, the attribution values represent direct marginal contributions to crop yield in physical units rather than transformed logarithmic scales.

Table 6. Global SHAP Feature Importance Ranking (Mean |SHAP| across 271 Test Observations)

Feature	Category	Mean |SHAP|	Contribution (%)
Plant Age (Years)	**Agronomic**	**0.454665**	**47.51%**
Land Area (Ha)	Operational	0.136992	14.32%
Annual Rainfall (mm)	Agroclimatic	0.104045	10.87%
Labor Ratio (Ha/Worker)	Operational	0.081391	8.51%
Estate Division	Management	0.081361	8.50%
Annual Rainy Days	Agroclimatic	0.052763	5.51%
Operational Disruption Index	Operational	0.031417	3.28%
Extreme Rain Months (>22 days)	Agroclimatic	0.009233	0.96%
Dry Months (<60mm)	Agroclimatic	0.005028	0.53%

Based on Table 6 and Figure 7, which illustrate the global SHAP feature importance ranking, Plant Age (Years) emerges as the most dominant factor, accounting for a contribution of 47.51%. This substantial contribution indicates that the biological age phase of the crop is the primary determinant of productivity variance at the block level. This finding aligns with previous empirical reports regarding the cumulative impact of crop age distribution on national oil palm yields in Malaysia [4]. Importantly, these SHAP attributions reflect predictive feature importance within XGBoost's empirical patterns rather than causal mechanisms, meaning that field interventions must combine these statistical insights with domain agronomic logic.

Furthermore, the Land Area (Ha) feature ranks second with a contribution of 14.32%. The magnitude of this feature's contribution indicates that the operational scale of the area plays a critical role in influencing inter-block productivity variations. This outcome aligns with prior studies indicating that land area, production scale, and workforce allocation are significantly associated with plantation productivity [10]. It should be clarified that land area itself does not directly govern biological crop yield; rather, it serves as an operational proxy reflecting block spatial scale, management intensity, and logistical harvesting efficiency. Larger contiguous block sizes optimize harvester routing and supervisory oversight, whereas fragmented or smaller blocks often suffer from higher operational downtime and inconsistent labor allocation. Subsequently, the Annual Rainfall (mm) variable ranks third at 10.87%, reinforcing the vital importance of adequate water supply to support fruit development [26].

Lastly, the Labor Ratio (Ha/Worker) feature occupies the fourth position, contributing 8.51% to the model. This finding provides empirical evidence that the proxy ratio of land area workload relative to the estimated number of harvesters exerts a significant influence on oil palm productivity [9].

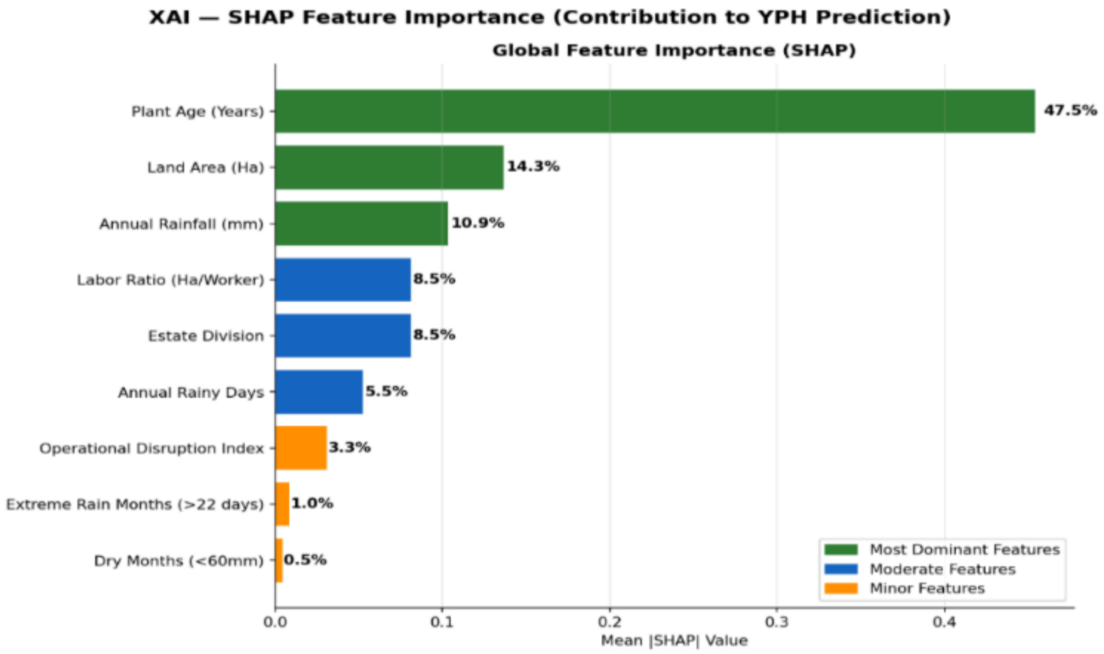

Figure 7. SHAP global feature importance: mean |SHAP| and percentage contribution of each feature

It should be clarified that the Estate Division feature, which contributes 8.50% to SHAP importance, reflects cross-sectional patterns across regions rather than a direct causal management effect. In field-level plantation data, division identifiers indirectly encompass unrecorded localized factors, such as variations in micro-soil conditions, land topography, historical road and drainage infrastructure, and localized work habits. Therefore, yield differences observed between divisions represent a composite of these unique physical and operational site characteristics, rather than purely isolated managerial performance.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

Conversely, threshold-based features such as Extreme Rain Months (0.96%) and Dry Months (0.53%) exhibit very minor SHAP contributions. This occurs primarily because overall annual climate patterns (Annual Rainfall and Annual Rainy Days) already capture the dominant climatic variance influencing crop yield. Furthermore, while extreme monthly weather events did occur during the 2018–2024 observation period, their overall frequency across the dataset was relatively low. Because these extreme monthly conditions were uncommon across total block-year observations, the XGBoost model relied more heavily on macro-climatic features to make its primary yield predictions.

Further analysis of Figure 8, successfully uncovers the direction of feature influences that typically remain undetected by standard correlation metrics. As an illustration, mature crops—those in the Prime phase—consistently yield significantly positive SHAP values, while younger crops conversely suppress the predictive outcomes. Furthermore, a high labor ratio induces negative SHAP values, signifying that an excessive land area workload allocated per harvester directly impairs harvesting intensity. This outcome demonstrates the utility of Explainable AI (XAI) approaches for visualizing complex feature interactions and non-linear patterns in agricultural datasets [27].

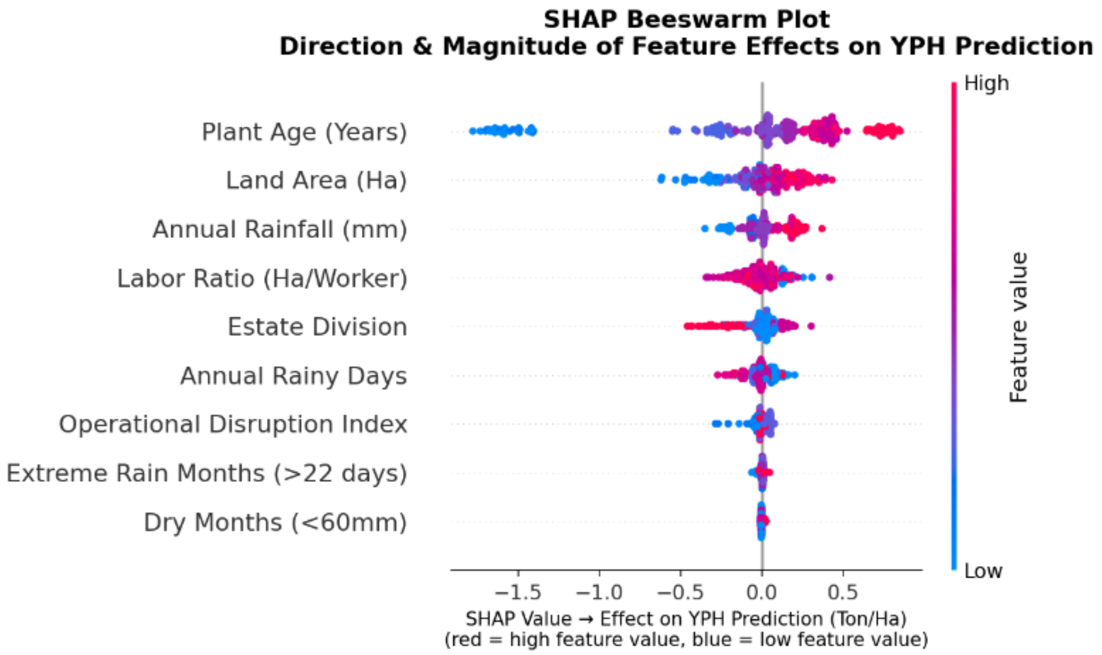

Figure 8. SHAP beeswarm plot: direction and magnitude of the impact of each feature on YPH prediction across 271 test observations

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

Analysis of the dependence plot (Figure 9) reveals the non-linear dynamics of the underlying model features. The relationship between plant age and YPH values forms a curve that increases during the initial stage and begins to plateau upon reaching 10–12 years of age. This pattern indicates that the marginal effect of increasing crop age on productivity predictions diminishes after entering this phase. This finding aligns with the crop productivity cycle concept, which demonstrates that productivity rises until the mature phase before gradually declining at an older age [4]. Conversely, the annual rainfall variable displays a positive influence trend on YPH predictions within the observed range of 4,000–6,300 mm. This outcome is consistent with previous studies reporting a positive relationship between rainfall and FFB production [26]. However, several studies have also indicated that excessive rainfall can exert negative impacts on both productivity and harvesting activities if it surpasses the optimum condition [5]. The positive SHAP attribution observed for high annual rainfall within the 4,000–6,300 mm range is context-specific to the local agroclimatic and infrastructural conditions of the studied estate in Kalimantan. In this region, high evapotranspiration rates and sandy-clay soil properties demand substantial continuous moisture to prevent yield-limiting water deficits. Furthermore, the estate's functional block-level drainage network mitigates prolonged root zone waterlogging during extreme wet periods. However, it is important to clarify that this trend is site-specific; in other geographical regions with poor drainage or flat lowlands, excessive rainfall exceeding 4,000 mm frequently reduces productivity due to root anoxia, severe nutrient leaching, and disrupted floral pollination.

Evaluation of the operational variables demonstrates non-linear patterns for the Land Area and Labor Ratio features. For the Land Area feature, the contribution toward YPH predictions tends to increase alongside expanding land size up to approximately 40 Ha. Meanwhile, the Labor Ratio feature shows negative SHAP values when the land workload per harvester falls within the 7–9 Ha range. This pattern indicates that as the land area allocated to each harvester expands, the model inherently predicts lower YPH productivity. However, it should be noted that the Labor Ratio feature is constructed as an operational proxy based on standard block area allocations rather than daily active worker attendance logs. Consequently, its SHAP attribution partially reflects block spatial scale alongside labor deployment density. These results demonstrate that the XGBoost framework coupled with Explainable AI (XAI) is capable of uncovering non-linear

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

relationships among variables that are otherwise difficult to represent using conventional linear regression approaches.

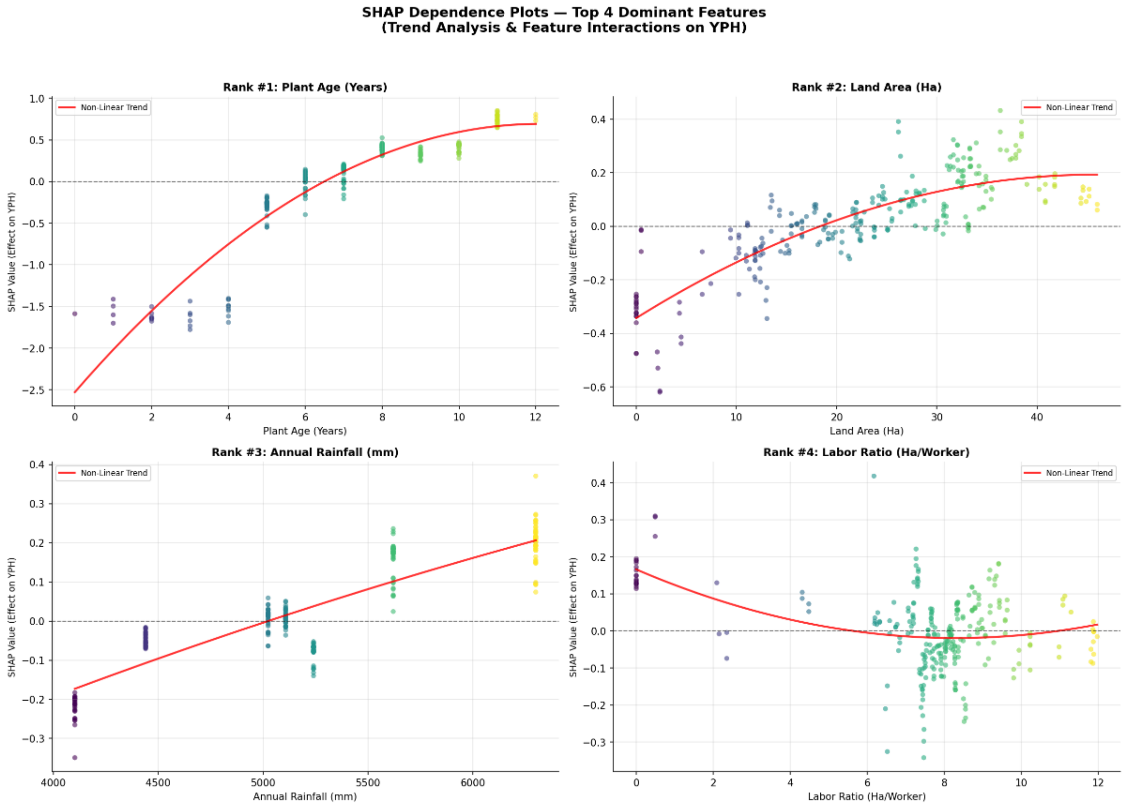

Figure 9. SHAP dependence plots for the top four dominant features: plant age, land area, annual rainfall, and labor ratio

To explicitly evaluate pairwise feature dependencies, SHAP interaction values were computed across all feature combinations on the test set (Figure 10). The top pairwise interactions were dominated by Plant Age interacting with Annual Rainy Days (Mean |SHAP Interaction| = 0.0351), Land Area (0.0247), Annual Rainfall (0.0246), and Estate Division (0.0205). These empirical interactions confirm that the biological crop maturity phase non-linearly modulates how agroclimatic stress and field operational scale suppress block-level yield.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

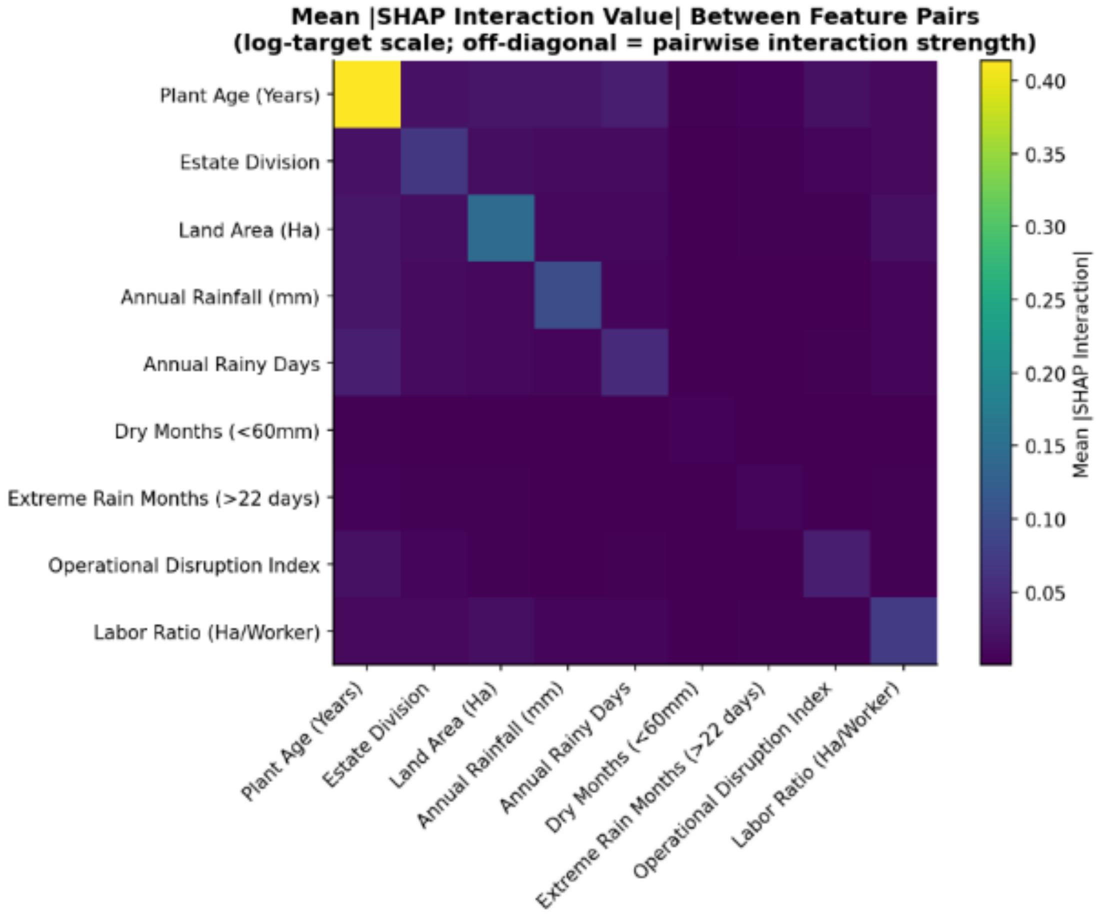

Figure 10. Heatmap matrix of mean absolute SHAP interaction values across feature pairs.

To evaluate the stability of feature importance attributions, a comparative validation was conducted against native XGBoost Gain-based importance and model-agnostic Permutation Importance (n = 20 iterations on testing set) as illustrated in Figure 11. Spearman rank correlation analysis revealed a strong alignment between SHAP feature attributions and Permutation Importance (ρ = 0.917, p < 0.001), with both methods consistently ranking Plant Age, Land Area, and Annual Rainfall as the top three yield drivers. Conversely, native Gain-based importance exhibited lower rank correlation with SHAP (ρ = 0.250) due to its inherent splitting-frequency bias. While this strong convergence between SHAP and Permutation Importance supports the empirical robustness of the attribution hierarchy, it should be noted that statistical alignment confirms structural stability rather than proving the complete absence of algorithmic bias.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

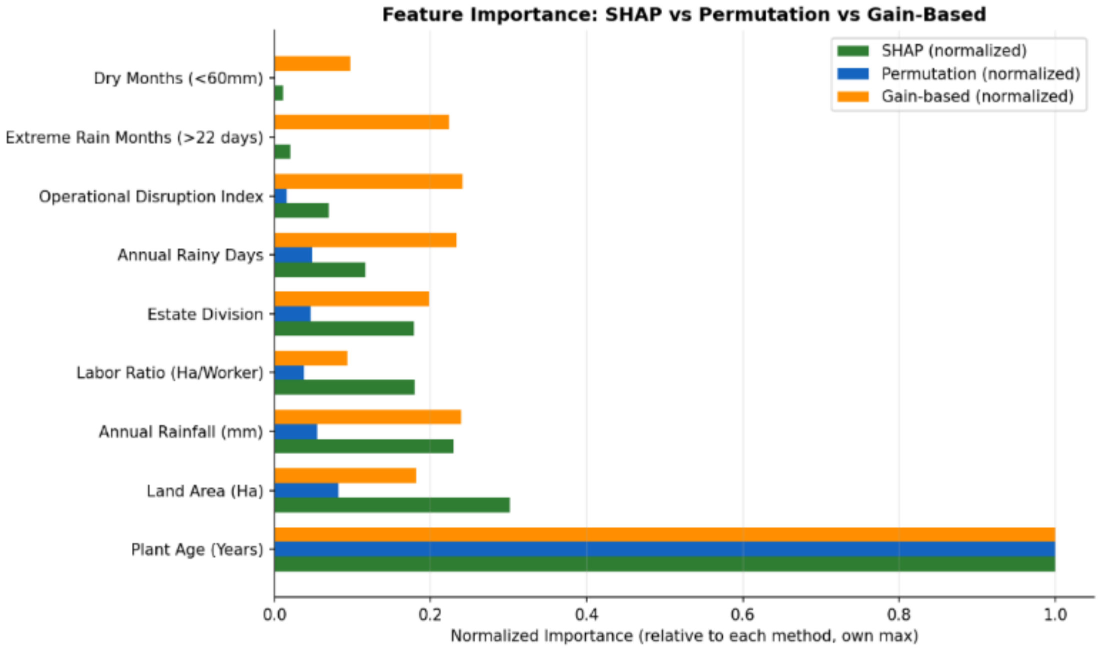

Figure 11. feature importance comparison across SHAP, Permutation Importance, and Gain-based XGBoost importance methods.

3.5. Managerial Insights Dashboard

To translate empirical XAI outputs into actionable operational strategies, an integrated Managerial Insights Dashboard was synthesized (Figure 12). The dashboard consolidates the global SHAP feature hierarchy, categorical contribution proportions, actual versus predicted yield trends, and data-driven strategic recommendations designed for estate-level decision support.

Based on the SHAP feature aggregation, the Agronomic category represents the most dominant driver of yield variability, accounting for 47.51% of total model importance. This is followed by the Operational category (26.11%), Agroclimatic factors (17.87%), and Management division-level influences (8.50%). Among individual features, Plant Age (Years) ranks as the primary determinant, followed by Land Area (Ha), Annual Rainfall (mm), Labor Ratio (Ha/Worker), and Estate Division. The close alignment between actual and predicted Yield Per Hectare (YPH) trends across test samples visually reinforces the model's stability in tracking field-level yield dynamics.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

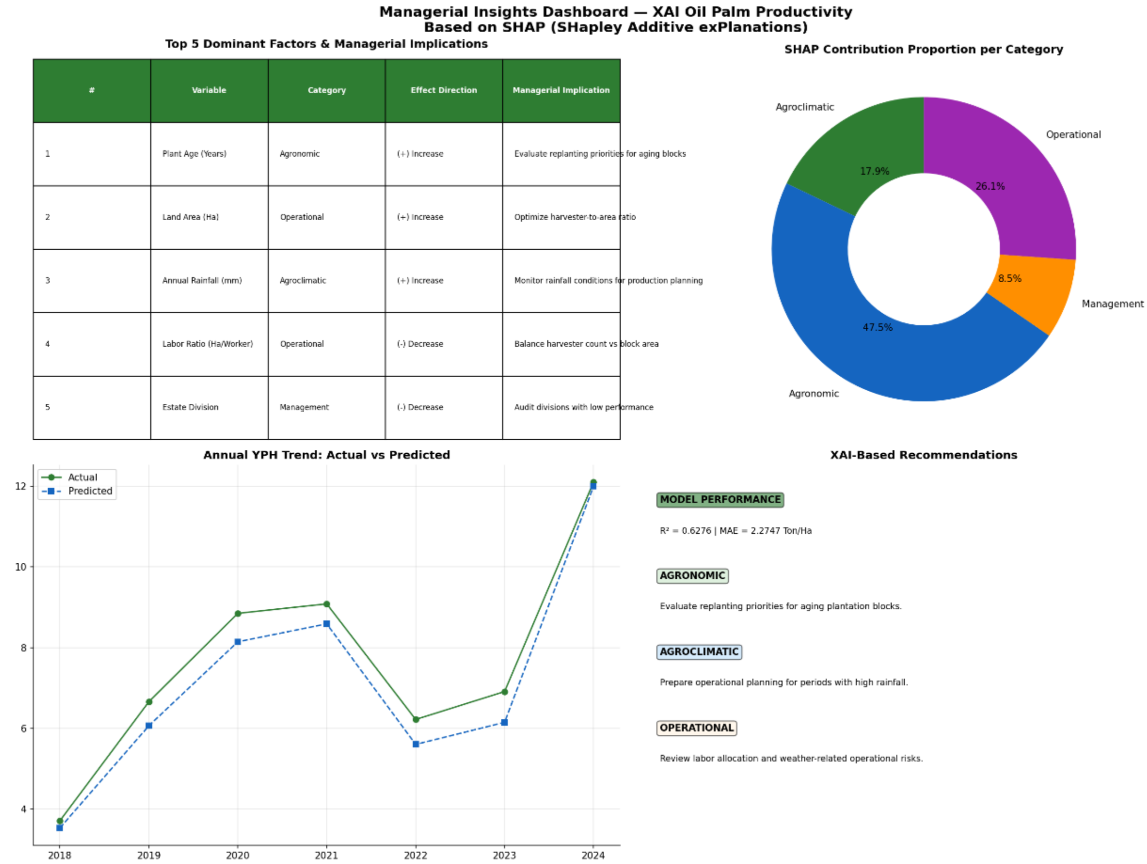

Figure 12. XAI managerial dashboard: SHAP contribution proportion per category, actual vs. predicted YPH trends, and strategic recommendations

Based on these transparent interpretability insights, key strategic directives are formulated to support decision-making for enhancing block productivity:

1) Agronomic Management Directives: Prioritize intensive maintenance, balanced fertilization, and targeted canopy management on mature blocks (TM-2 and TM-3) to sustain peak yield potential.
2) Operational Efficiency: Maintain the labor workload ratio within the optimal threshold of 7–9 Ha per worker to prevent harvesting delays and under-recovery of fresh fruit bunches (FFB).
3) Agroclimatic Resilience: Implement proactive drainage infrastructure in flood-prone blocks during extreme rainfall months (> 22 rainy days) to minimize operational field disruptions.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

However, it is critical to frame these XAI-derived managerial thresholds as data-driven decision-support heuristics rather than rigid or autonomous prescribing rules. While the XGBoost model accurately captures complex historical non-linear patterns, practical field deployment necessitates close validation with domain experts and plantation managers. Resident agronomists and estate managers must contextualize these empirical recommendations against unobserved, real-time field variables—such as daily labor availability, localized micro-topography, and immediate weather fluctuations. Future operational rollouts should incorporate structured Focus Group Discussions (FGDs) with estate management to systematically validate and calibrate model outputs prior to full-scale implementation.

3.6. Practical Implications and Research Limitations

While the proposed XGBoost-SHAP framework provides valuable insights into non-linear factor interactions, several methodological and operational caveats must be acknowledged when interpreting the results:

1) Model Generalization and Validation Constraints: The model's predictive capability demonstrated a significant performance reduction when evaluated under Grouped Split (R^2 = 0.2300) and Time-based Split (R^2 = -0.1106) compared to the baseline Random Split (R^2 = 0.6276). This discrepancy demonstrates that random block-year splitting carries inherent spatial and temporal autocorrelation risks, potentially overestimating model performance due to identical physical blocks recurring across train-test partitions. The model captures localized cross-sectional variance effectively but requires continuous multi-year recalibration before deploying across unseen geographic blocks or distinct climatic phases.
2) Interpretability versus Causal Inference: Feature attributions derived via SHAP must be interpreted strictly as post-hoc model explanations rather than empirical causal proofs. While SHAP accurately identifies how features contribute to the model's conditional expectations, it cannot establish direct biochemical or counterfactual causation.
3) Operational Proxies and Latent Confounders: Several engineered features reflect operational and spatial proxies rather than direct agronomic drivers. Specifically, *Land Area* captures block-level management intensity, boundary drainage efficiency, and operational scale heterogeneity rather than intrinsic

palm physiology. Similarly, *Estate Division* acts as a composite proxy capturing unobserved spatial confounders, including localized soil series variations, infrastructure quality, logistical distance, and supervisory management practices. Furthermore, the *Labor Ratio* metric is structurally tied to the operational baseline assumption of 8 ha per harvester, meaning its interaction effects reflect staffing allocation constraints relative to block size rather than isolated labor productivity.

4) Regional Agroclimatic Specificity: The positive yield response observed within the high precipitation band of 4,000–6,300 mm is contextually specific to the equatorial, high-rainfall environment of Kalimantan. This regional hydrological balance cannot be extrapolated as a universal agronomic optimum for oil palm cultivation in drier climatic zones with different soil water retention capacities.
5) Framework Readiness and Managerial Decision Support: Consequently, this framework serves as an exploratory decision-support foundation rather than an immediately field-ready prescriptive tool. Practical field deployment requires operational recalibration and participatory validation through Focus Group Discussions (FGD) and trial implementations with estate managers to align model recommendations with field logistics and agronomic standard operating procedures.

4. CONCLUSION

This study applies a 9-feature multivariate XGBoost-SHAP framework—integrating agronomic, agroclimatic, and operational domains—to advance block-level oil palm productivity analysis, delivering core contributions across three primary dimensions: (1) prediction performance, achieving an R^2 of 0.6276, MAE of 2.2747 Tons/Ha, RMSE of 3.3303 Tons/Ha, and an MAPE of 28.73% on the baseline random hold-out split; (2) explainability findings, revealing through SHAP interpretation that plant age is the dominant yield driver (47.51% contribution), followed by land area (14.32%), annual rainfall (10.87%), and the harvester workload ratio (8.51%); and (3) managerial implications, synthesizing empirical insights into an interpretive heuristic prototype dashboard. However, evaluating the model under grouped block (R^2 = 0.2300) and time-based (R^2 = -0.1106) validation schemes revealed much weaker performance, indicating that the baseline random split performance is likely optimistic. Consequently, the framework currently supports retrospective block-level interpretation more strongly than future yield forecasting. Furthermore, several empirical limitations must be acknowledged: the findings are restricted to a single plantation in Central Kalimantan without external validation, subject to block-year leakage risks, rely on an area-based labor proxy, and lack granular data on fertilization, harvesting rotation logs, and soil attributes. To address these constraints, future work must evaluate model generalizability across external estates, incorporate direct daily labor records, detailed fertilizer logs, harvesting rotation frequencies, and soil variables, while systematically validating the dashboard heuristics through structured Focus Group Discussions (FGDs) with plantation managers before field implementation.

ACKNOWLEDGMENT

The authors thank the management and staff of the oil palm plantation in Central Kalimantan for providing the historical research data, as well as the Department of Informatics Engineering at Universitas Muria Kudus for institutional and academic support throughout this study.

REFERENCES

[1] L. S. Woittiez, M. Slingerland, M. van Noordwijk, A. J. Silalahi, J. van Heerwaarden, and K. E. Giller, “People, Palms, and Productivity: Testing Better Management Practices in Indonesian Smallholder Oil Palm Plantations,” *Agriculture*, vol. 14, no. 9, p. 1626, Sep. 2024, doi: 10.3390/agriculture14091626.

[2] N. N. Kamil, S. Xiao, S. N. Syed Salleh, H. Xu, and C. C. Zhuang, “Nonlinear impacts of climate anomalies on oil palm productivity,” *Heliyon*, vol. 10, no. 15, p. e35798, Aug. 2024, doi: 10.1016/j.heliyon.2024.e35798.

[3] S. Fleiss, C. J. McClean, H. King, and J. K. Hill, “Limited impacts of climatic conditions on commercial oil palm yields in Malaysian plantations,” *CABI Agriculture and Bioscience*, vol. 3, no. 1, p. 59, Sep. 2022, doi: 10.1186/s43170-022-00127-1.

[4] J. F. Khor *et al.*, “Impact Comparison of El Niño and Ageing Crops on Malaysian Oil Palm Yield,” *Plants*, vol. 12, no. 3, p. 424, Jan. 2023, doi: 10.3390/plants12030424.

[5] M. J. Suhartono, M. T. S. Budiastuti, M. Mujiyo, and S. Gunawan, “The Effect of Climate Anomalies on Oil Palm Productivity: A Quantitative Analysis,” *BIO Web Conf.*, vol. 189, p. 01008, Oct. 2025, doi: 10.1051/bioconf/202518901008.

[6] L. S. Woittiez, M. T. van Wijk, M. Slingerland, M. van Noordwijk, and K. E. Giller, "Yield gaps in oil palm: A quantitative review of contributing factors," *European Journal of Agronomy*, vol. 83, pp. 57–77, Feb. 2017, doi: 10.1016/j.eja.2016.11.002.

[7] J. P. Monzon *et al*., "Agronomy explains large yield gaps in smallholder oil palm fields," *Agric. Syst*., vol. 210, p. 103689, Aug. 2023, doi: 10.1016/j.agsy.2023.103689.

[8] Y. L. Lim *et al*., "Too little, too imbalanced: Nutrient supply in smallholder oil palm fields in Indonesia," *Agric. Syst*., vol. 210, p. 103729, Aug. 2023, doi: 10.1016/j.agsy.2023.103729.

[9] J. Ginting and L. Lindawati, "Analisis Faktor-Faktor yang Mempengaruhi Produktivitas Tenaga Kerja Pemanen Kelapa Sawit (Elaeis Gunieensis Jacq.) PT. Pekebunan Nusantara IV Regional II," *Jurnal Ekonomi Pertanian dan Agribisnis*, vol. 9, no. 3, pp. 891–898, Aug. 2025, doi: 10.21776/ub.jepa.2025.009.03.3.

[10] H. Wikantyasning and Marseto, "The Effect Of Area Area, Production, and Number Of Workers On The Productivity Of Oil Palm Plantations In Indonesia In 2014-2023," *Oikonomia : Journal of Management Economics and Accounting*, vol. 2, no. 3, pp. 45–56, May 2025, doi: 10.61942/oikonomia.v2i3.340.

[11] W. Hasanah, N. Susanti, and D. Agustina, "The Effect of Workload on Work Productivity Oil Palm Harvester," *Indonesian Journal of Global Health Research*, vol. 6, no. 3, pp. 1699–1710, Jun. 2024, doi: 10.37287/ijghr.v6i3.3368.

[12] L. Nurfalah, A. A. Dwiyahreni, N. L. Winarni, K. Mizuno, and P. Grassini, "Preliminary Results of Relationship between Oil Palm Harvest Losses and Harvest Interval in Riau and West Kalimantan, Indonesia," *Jurnal Penelitian Pendidikan IPA*, vol. 9, no. 12, pp. 10721–10726, Dec. 2023, doi: 10.29303/jppipa.v9i12.6120.

[13] R. E. de Vos *et al*., "Shortening harvest interval, reaping benefits? A study on harvest practices in oil palm smallholder farming systems in Indonesia," *Agric. Syst*., vol. 211, p. 103753, Oct. 2023, doi: 10.1016/j.agsy.2023.103753.

[14] F. N. Mohd Nain, N. H. Ahamed Hassain Malim, R. Abdullah, M. F. Abdul Rahim, M. A. Ahmad Mokhtar, and N. S. Mohamad Fauzi, "A Review of an Artificial Intelligence Framework for Identifying the Most Effective Palm Oil Prediction," *Algorithms*, vol. 15, no. 6, p. 218, Jun. 2022, doi: 10.3390/a15060218.

[15] G. N. Elwirehardja, T. Suparyanto, Miftakhurrokhmat, and B. Pardamean, "Determining Variables Associated with Annual Oil Palm Yield: An Explainable Gradient Boosting Approach," *Procedia Comput. Sci*., vol. 227, pp. 262–271, 2023, doi: 10.1016/j.procs.2023.10.524.

[16] N. Khan *et al.*, "Prediction of Oil Palm Yield Using Machine Learning in the Perspective of Fluctuating Weather and Soil Moisture Conditions: Evaluation of a Generic Workflow," *Plants*, vol. 11, no. 13, p. 1697, Jun. 2022, doi: 10.3390/plants11131697.

[17] F. Watson-Hernández, N. Gómez-Calderón, and R. P. da Silva, "Oil Palm Yield Estimation Based on Vegetation and Humidity Indices Generated from Satellite Images and Machine Learning Techniques," *AgriEngineering*, vol. 4, no. 1, pp. 279–291, Mar. 2022, doi: 10.3390/agriengineering4010019.

[18] Y. Ang *et al.*, "Block-scale Oil Palm Yield Prediction Using Machine Learning Approaches Based on Landsat and MODIS Satellite Data," *Journal of Advanced Research in Applied Sciences and Engineering Technology*, vol. 45, no. 1, pp. 90–107, Mar. 2025, doi: 10.37934/araset.45.1.90107.

[19] A. M. Shimaoka, R. C. Ferreira, and A. Goldman, "The evolution of CRISP-DM for Data Science: Methods, Processes and Frameworks," *SBC Reviews on Computer Science*, vol. 4, no. 1, pp. 28–43, Oct. 2024, doi: 10.5753/reviews.2024.3757.

[20] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA: Association for Computing Machinery, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[21] D. R. Roberts *et al.*, "Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure," *Ecography*, vol. 40, no. 8, pp. 913–929, Aug. 2017, doi: 10.1111/ecog.02881.

[22] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, Long Beach, CA, USA: Curran Associates, Inc., 2017, pp. 4768–4777.

[23] M. Li, H. Sun, Y. Huang, and H. Chen, "Shapley value: from cooperative game to explainable artificial intelligence," *Autonomous Intelligent Systems*, vol. 4, no. 1, p. 2, Feb. 2024, doi: 10.1007/s43684-023-00060-8.

[24] R. N. V. J. Mohan, P. S. Rayanoothala, and R. P. Sree, "Next-gen agriculture: integrating AI and XAI for precision crop yield predictions," *Front. Plant Sci.*, vol. 15, p. 1451607, Jan. 2025, doi: 10.3389/fpls.2024.1451607.

[25] P. Oettli, S. K. Behera, and T. Yamagata, "Climate Based Predictability of Oil Palm Tree Yield in Malaysia," *Sci. Rep.*, vol. 8, no. 1, p. 2271, Feb. 2018, doi: 10.1038/s41598-018-20298-0.

ISI Journal of Information Systems and Informatics
Published By Asosiasi Doktor Sistem Informasi Indonesia

[26] W. Atmaka Aji, S. Aryo Saloko, I. Kurniawan, and M. Ahmad Nurcahya, "The Effect of Rainfall and Rain Days On Palm Oil Production," *Tropical Plantation Journal*, vol. 1, no. 1, pp. 16–20, Jan. 2022, doi: 10.56125/tpj.v1i1.3.

[27] D. G. Pai, M. Balachandra, and R. Kamath, "Explainable AI in agriculture: review of applications, methodologies, and future directions," *Engineering Research Express*, vol. 7, no. 3, p. 032202, Sep. 2025, doi: 10.1088/2631-8695/ae086a.