ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may2119

An Ensemble Machine Learning Approach for Pre-Harvest Rice Yield Forecasting in Sierra Leone Using Sentinel-2 and Multi-Source Climate Data

Abdul Koroma[1]

[1]Nankai University, Tianjin, China

Publication Date: 2026/06/17

Abstract: The current research focuses on studying rice production, which accounts for about 62% of the national energy intake, as an important element of food security in Sierra Leone. Currently, there is no objective approach to predicting rice yields in a timely manner. In this work, we propose a machine learning model that uses multi-temporal Sentinel-2 images, the ERA5-Land and CHIRPS weather datasets, and climate forecasting datasets to predict early-season rice yields for Kambia and Bombali districts in Sierra Leone. A training dataset with 1,774 samples of bi-weekly data collected during seven rice-growing seasons (2018-2024) is used to optimize an ensemble consisting of Ridge Regression, Random Forest, and XGBoost. The resulting ensemble demonstrates excellent performance on an independently selected testing dataset (2023-2024): R^2 = 0.963, RMSE = 16.87 kg/ha, and MAPE = 0.50%, significantly outperforming the existing benchmarks. Temporal trend-based features predicting an increase in rice yield by 11.5% become key predictors (account for 42% of total importance), along with vegetative index-based features and cumulative Growing Degree Days. Predictions for the early season of rice production for 2025 illustrate the stability and accuracy of prediction models until five months before harvesting.

Keywords: *Rice Yield Prediction; Sentinel-2; Machine Learning; Ensemble Model; Sierra Leone; Food Security; Early Warning Systems.*

How to Cite: Abdul Koroma (2026) An Ensemble Machine Learning Approach for Pre-Harvest Rice Yield Forecasting in Sierra Leone Using Sentinel-2 and Multi-Source Climate Data. *International Journal of Innovative Science and Research Technology*, 11(5), 4418-4429. https://doi.org/10.38124/ijisrt/26may2119

I. INTRODUCTION

Rice (*Oryza sativa* L.) is the core foundation of the national food system in Sierra Leone. As the popular Krio saying goes, “If I have not eaten rice, I have not eaten anything,” pointing to the nutritional importance of the crop, where rice is not simply one of the staples but rather defines the very notion of a meal, making up 62% of the total caloric intake of the country per day and costing some 25% of the budget of households across the country [24]. In that respect, this saying encapsulates both the nutritional importance and cultural significance of rice. At the same time, the livelihood of more than 70% of the rural population is supported by the crop. With such a high level of dependence, the level of imports becomes especially significant: at present, Sierra Leone imports 40% of its total rice consumption annually [21, 10]. This vulnerability has already manifested itself during the food price crisis of 2007-2008 and during the pandemic period [8]. Such is the problem addressed through the goal of 75% self-sufficiency by 2027 in the National Agricultural Development Plan [22]; however, reaching that goal will be possible only with planning and an understanding of the quantity of production.

Kambia and Bombali, being two of the three rice-producing districts within the Northern Province of Sierra Leone, provide the concrete context required to answer the question in this paper. The two are part of the main producing region in the country, having rainfall patterns ranging between 2,200-2,800 mm [19] annually in alluvial and lowlands suitable for mangrove swamp, inland valley, and boliland rice. They enjoy fertile soils and existing agricultural infrastructure in the area. The important institution in Kambia is the Rokupr Agricultural Research Centre (RARC), where, since its establishment in 1934, most of the improved rice varieties currently available have been developed [23]. Yield at the district level in both areas has seen a notable increase, going up from 2,400 kg/ha in 2018 to projections of 3,000 kg/ha by 2024 [25, 8, 9]. Increases of 57-64 kg/ha each year have been brought about by the adoption of improved varieties, access to fertilizers, and improved extension programs. The aforementioned increases are not just

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may2119

contextual elements in this paper but the main prediction in the dataset, whereby depending on how well a model handles that prediction defines its success or failure in making predictions for future seasons.

There does not seem to be any concrete monitoring mechanism for this trend currently. The yield information provided by the government is derived mainly from inspections by extension officers and household surveys, making it very difficult to predict any shortage ahead of time [20, 15]. The process of obtaining estimates from these sources is lengthy and prone to observer error accumulation. Moreover, once a shortage has been noted, the response window has often passed as the import contracts take time [12]; humanitarian response requires earlier probabilistic information, and traders need to be able to make accurate estimates of the supply available.

Two factors have made it more viable to adopt a better strategy. This new strategy can be seen in earth observation. The Sentinel-2 satellite launched by the European Space Agency provides a five-day revisiting period of global land coverage at ten meters' resolution for free [17, 18]. The continuous recording of crops' condition through an entire season is possible without making trips to the field. This, when combined with machine learning ensemble methods, becomes the basis for yield forecasting months ahead of time before harvesting. Specifically, machine learning ensemble models like gradient boosting and bagging have reached a point where they are the best methods that are available nowadays for making predictions about the yield from agriculture due to their ability to model complex interactions between vegetative condition, climatic pressure, and yield [6, 16].

There has been no previous study in the literature that used such a combination of tools for predicting rice production in Sierra Leone. This paper presents the first systematic methodology that combines Sentinel-2 satellite imagery, multi-source climatic datasets (ERA5-Land and CHIRPS), seasonal climatic forecasts, and machine learning algorithms to predict rice yields in Sierra Leone, offering high-accuracy annual forecasts and consistent pre-season forecasts about five months before harvest. Two important results stand out from the analysis. The first is related to the model structure; the trend of increasing yields through the years should be treated explicitly as a predictor. Without doing so, the well-defined ensemble model fails to forecast out-of-sample observations, as indicated by the negative R^2 [14]. This insufficiency is not due to calibration issues, but is an inherent structural insufficiency. A predictive model based on yield observations from 2018 through 2022 lacks a basis for predicting the high levels of yield seen in 2023 through 2024 unless it takes into account the underlying trend, which must be incorporated into the feature set. This is demonstrated explicitly, along with the exact processes causing this deficiency, and we show how the inclusion of trends over time helps resolve the issue, producing an ensemble R^2 score of 0.963 and MAPE of 0.50% when applied to unseen data. Any research team working on predictive models of yield in developing agricultural environments, where yields are increasing, such as most of Sub-Saharan Africa, would find this their key take-home message. The second insight is about practicality. By replacing unknown months with calibrated seasonal climate predictions [13, 12], the trained model provides district-wise estimates of crop harvests up to five months ahead of time in dry, normal, and wet situations, along with uncertainties. The entire workflow runs on data from Sentinel-2, ERA5-Land, CHIRPS, Google Earth Engine, and Python [11]. Open, free, and operational under the limitations of an institutional environment like a national agriculture ministry that does not have the resources to build infrastructure.

II. LITERATURE REVIEW

- *The Agricultural Monitoring Gap in West Africa*

There is little literature on West African studies that utilize machine learning to predict rice yield [1]. The value of R^2 of 0.72 was obtained in Nigeria through the use of MODIS satellite data, limited by the large pixel size (250 m-1km), causing a strong mixed pixels effect for regions having smallholder agriculture [2]. Random forest models were applied within several countries in West Africa, resulting in an R^2 of 0.68. Similarly, a similar analysis conducted in the Senegal River Valley region resulted in an R^2 of 0.78 and MAPE of 6.4%. This evidence confirms the viability of satellite forecasting but also draws attention to unsolved problems, namely, the impact of clouds on data collected from satellites, low resolution of spatial data, and the inability to deal with yield trends in fast-changing agricultural environments.

- *Machine Learning for Yield Prediction*

Ensemble methods, particularly Random Forest [3], Gradient Boosting [4], and XGBoost [5], have consistently produced state-of-the-art results across rice-growing regions. [6] According to a survey of fifty peer-reviewed articles, ensemble learning algorithms surpass single-algorithm-based models in over 80% of instances. An issue that keeps cropping up relates to nonstationary, whereby several models assume stationarity of predictor-yield relationships, leading to under-prediction of yields within systems undergoing constant innovation. This paper tackles this problem through the inclusion of the yield trend trajectory as a predictive variable.

- *Sentinel-2 for Rice Monitoring*

The 13 bands of the Multi-Spectral Instrument aboard Sentinel-2 feature four bands with 10 m resolution and six bands with 20 m resolution, three of which are red-edge bands highly responsive to chlorophyll. These arrangements allow the full tracking of phenology through the entire growing season. The use of Google Earth Engine (GEE) has made the analysis of Sentinel-2 data possible even for researchers from countries lacking access to these data. No studies have yet attempted to combine Sentinel-2 and multiple sources of climate data to forecast rice yields in Sierra Leone.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may2119

III. MATERIALS AND METHODS

➢ *Study Area*

The study covers Kambia and Bombali Districts in Sierra Leone's Northern Province (~11,093 km²; latitudes 8°45'N–9°45'N). The two divisions share the tropical monsoon type of climate, whereby rainfall prevails from May to November (2,200-2,800mm), coinciding with the entire planting period for rice. The district of Kambia (3,108 sq. Km.) has the Rokupr Agricultural Research Centre (RARC), which is the leading organization for rice research in Sierra Leone. The area of Bombali (7,985 sq. km.) has varied topography, including Makeni, the economic hub of the Northern Province. Rice growing takes place in four agro-ecological zones: Upland Rainfed (1.0-2.0 t/ha), Inland Valley Swamps (2.0-3.5 t/ha), Riverine Flood Plains (2.5-4.0 t/ha).

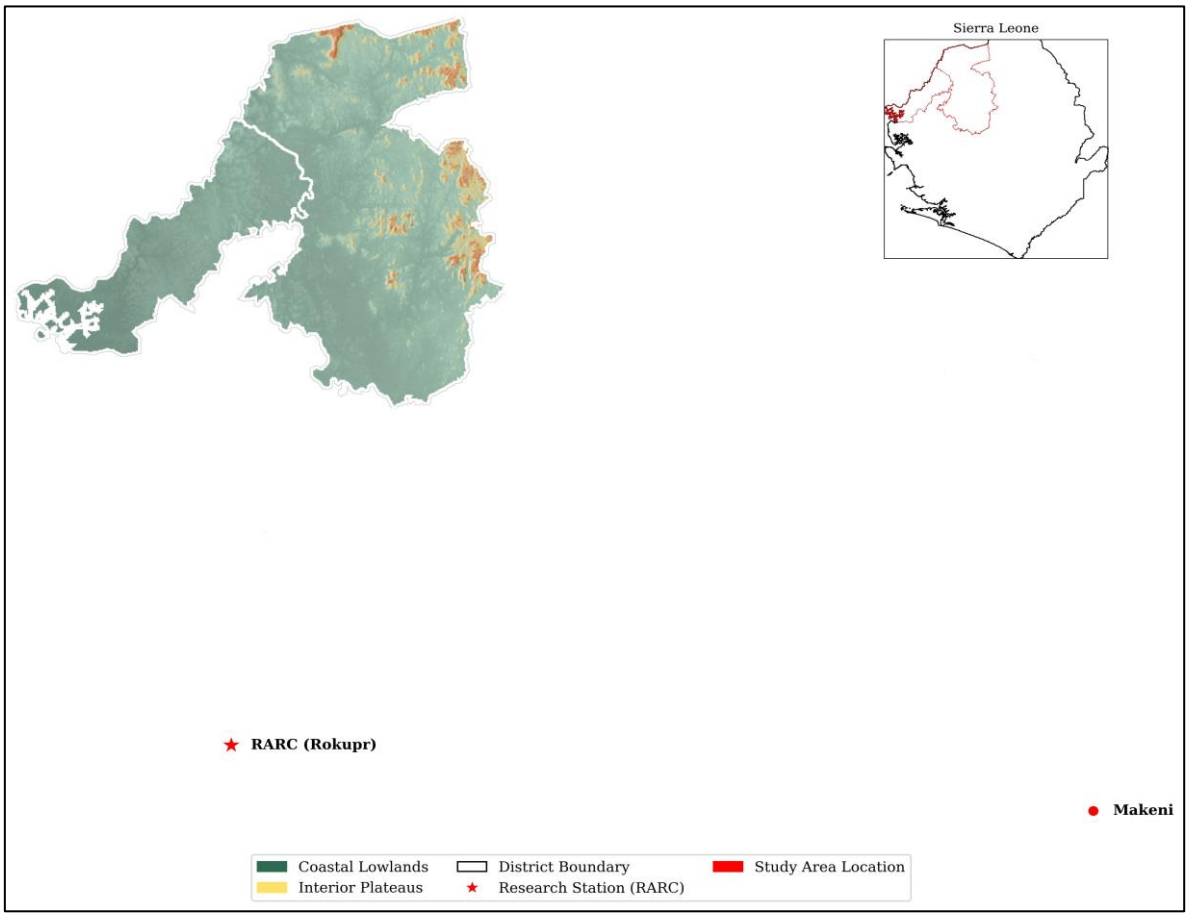

Fig 1 Study Area Kambia and Bombali Districts, Sierra Leone

The main map displays the elevation gradient from the low-lying coastal plains of Kambia to the interior plateaus of Bombali, with white lines delineating district boundaries. Key institutional and administrative hubs, including the Rokupr Agricultural Research Centre (RARC) and the city of Makeni, are indicated with symbols. The top-right inset provides the national context, showing the location of the study districts within Sierra Leone.

➢ *Research Design and Methodological Framework*

The framework follows a six-stage sequential pipeline that transforms raw satellite imagery and climate records into operationally useful yield forecasts (Figure 2) [11, 6]. Phase I entails the collection and pre-processing of multi-temporal imagery from Sentinel-2 satellite sensors as well as the respective gridded climate variables for each crop-growing season, ranging from 2018 to 2024. Phase II comprises aggregating official data on yield per district and assigning each data point to its bi-weekly satellite measurements per year and district. Phase III is about developing 28 predictors split into four groups, namely vegetation indices, climate factors, agronomic indicators, and temporal-spatial terms, which cover within-season crop dynamics and long-term trends in crop productivity. Phase IV consists of training and validating three machine learning algorithms and building a weighted model ensemble based on year-blocked five-fold cross-validation on training years only (i.e., 2018-2022). Phase V assesses model performance using a completely separate test set (2023-2024) unseen by all previous phases. Finally, Phase VI replaces unobserved months in the future with seasonal climate forecasts and generates crop yield forecasts for early-season assessment [12, 13].

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may2119

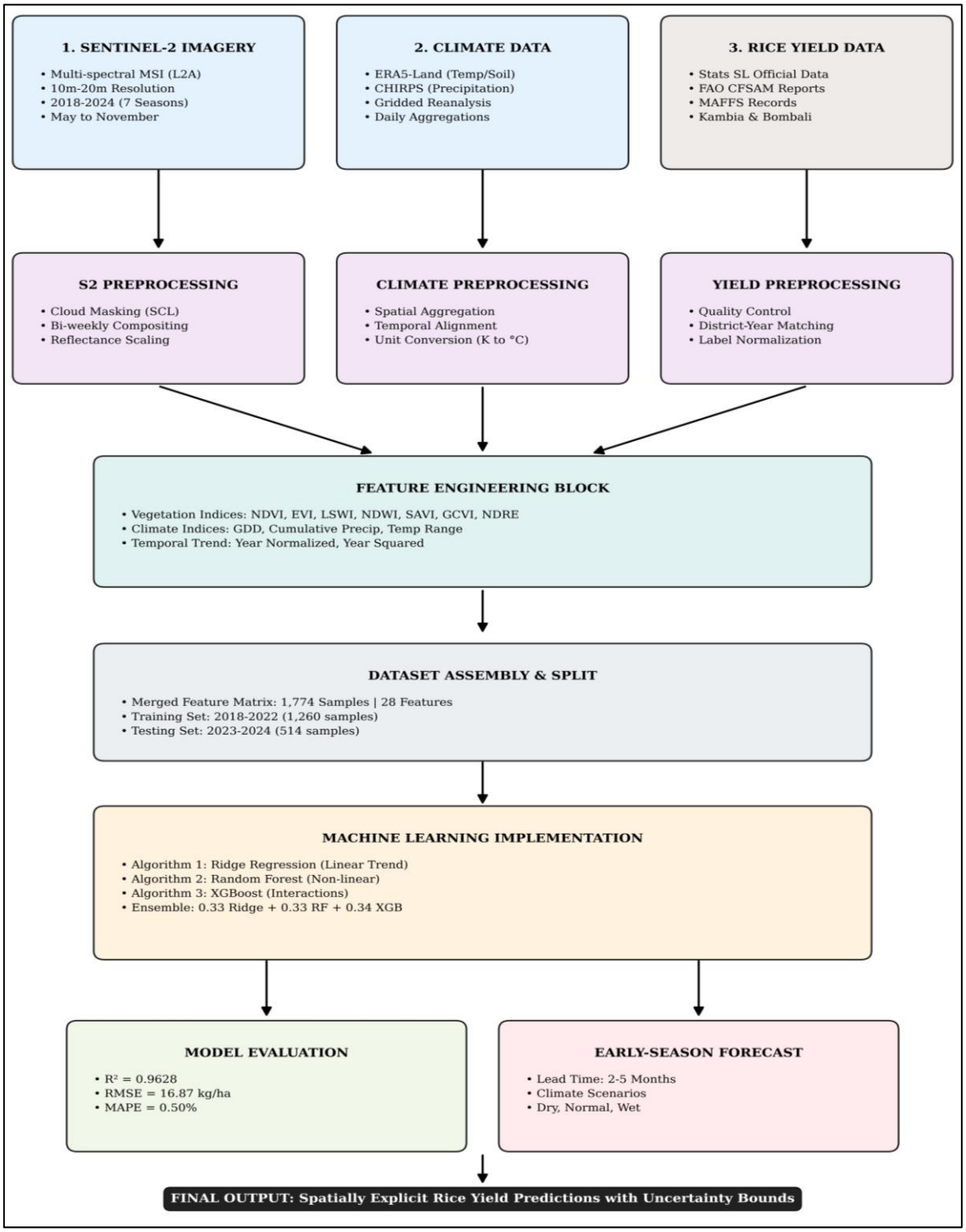

Fig 2 Research Framework Flowchart Illustrating the Six-Stage Analytical Pipeline

There are some issues that cut across all phases and therefore need to be addressed directly. The first thing is that the train test separation must always adhere strictly to the time aspect. This means that the model is always tested against future data rather than past data, as would be the case in random cross-validation [6]. Second, both the standardization of the features and fitting of the models happen solely using the training set data; statistics calculated on the training data set are used to transform test data set observations to prevent information leakage. Third, the use of time trends as input variables rather than detrending the output variable was an intentional decision based on the realization that improvements in yields in both areas can be predicted [14]. Fourth, the use of an eight-grid-cell sample for each district ensures that the effective size of the training dataset increases without altering the structure of the yield labels in each district. This technique ensures that the supervised learning problem is constrained by the resolution of the actual data available.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may2119

➢ *Data Sources*

Sentinel-2 Level-2A (Surface Reflectance) imagery was accessed via GEE (COPERNICUS/S2_SR_HARMONIZED) for May 1–November 30, 2018–2024. Cloud masking used the Scene Classification Layer (SCL), excluding cloud shadows (SCL=3), medium probability clouds (SCL=8), high-probability clouds (SCL=9), and thin cirrus (SCL=10). Biweekly (15-day) median composites generated up to 14 observation periods per season. Climate data came from ERA5-Land Daily Aggregated (~9 km) for temperature, soil moisture, solar radiation, and evapotranspiration, and CHIRPS Daily (~5 km) for precipitation. District-level yield data were compiled from [7], [8], [9], and FAOSTAT-based estimates [10].

Table 1 District-Level Rice Yield Dataset (2018–2024). Training: 2018–2022; Test: 2023–2024.

Year	District	Yield (kg/ha)	Source	Confidence
2018	Kambia	2,620	FAOSTAT + Regional Adj.	Medium
2018	Bombali	2,400	FAOSTAT + Regional Adj.	Medium
2019	Kambia	2,690	FAOSTAT + Regional Adj.	Medium
2019	Bombali	2,480	FAOSTAT + Regional Adj.	Medium
2020	Kambia	2,723	Statistics Sierra Leone	High
2020	Bombali	2,464	Statistics Sierra Leone	High
2021	Kambia	2,650	FAO CFSAM	High
2021	Bombali	2,550	FAO CFSAM	High
2022	Kambia	2,900	MAFFS Rice Strategy	High
2022	Bombali	2,700	MAFFS Rice Strategy	High
2023	Kambia	2,950	Trend Interpolation	Medium
2023	Bombali	2,780	Trend Interpolation	Medium
2024	Kambia	3,020	Projection	Low
2024	Bombali	2,850	Projection	Low

➢ *Vegetation Indices*

Seven vegetation indices were generated from the Sentinel-2 spectral bands in each bi-weekly composite image. Band notation: B2 = Blue, B3 = Green, B4 = Red, B5 = Red Edge, B8 = NIR, B11 = SWIR.

- *Normalized Difference Vegetation Index (NDVI)*

$$NDVI = \frac{B8 - B4}{B8 + B4} \quad (1)$$

NDVI acts as an indicator of vegetation cover. For healthy rice crops, NDVI ranges normally between 0.45 in the early stages of crop growth to 0.70 during the reproductive stage. NDVI has the tendency to become saturated in dense canopies, hence the need for additional indices.

- *Enhanced Vegetation Index (EVI)*

$$EVI = 2.5 \times \frac{B8 - B4}{B8 + 6 \cdot B4 - 7.5 \cdot B2 + 1} \quad (2)$$

The Enhanced Vegetation Index (EVI), which uses the blue band, addresses atmospheric aerosol influences and prevents saturation in dense canopies; thus, the index is very appropriate for assessing the productivity of the lowland rice growing areas of the study area.

- *Land Surface Water Index (LSWI)*

$$LSWI = \frac{B8 - B11}{B8 + B11} \quad (3)$$

The Land Surface Water Index (LSWI) is responsive to changes in the water content of leaves and paddy fields under flooded conditions, making LSWI an important indicator of flooded paddy during transplantation as well as water stress in the reproduction phase.

- *Normalized Difference Water Index (NDWI)*

$$NDWI = \frac{B3 - B8}{B3 + B8} \quad (4)$$

NDWI is used for detection of open water and crop moisture stress. Temporal change in NDWI during the cropping season indicates changes in the water content of plants due to changes in canopy growth.

- *Soil Adjusted Vegetation Index (SAVI)*

$$SAVI = \frac{B8 - B4}{B8 + B4 + L} \times (1 + L) \quad (5)$$

SAVI minimizes the effect of the reflectance of the soil in the early stages of growth of the canopy. L = 0.5 is the standard for the soil reflectance correction factor.

- *Green Chlorophyll Vegetation Index (GCVI)*

$$GCVI = \frac{B8}{B3} - 1 \quad (6)$$

The GCVI is responsive to chlorophyll concentration and green leaf area, showing linearity in response at a wider range of leaf area index than NDVI does. This property

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may2119

improves its usefulness as a measure of the nitrogen condition of crops during their growth cycle.

- *Normalized Difference Red Edge (NDRE)*

$$NDRE = \frac{B8 - B5}{B8 + B5} \quad (7)$$

The NDRE index takes advantage of the sensitivity of the red-edge band of Sentinel-2 (B5, 705 nm) to chlorophyll, even in situations with high biomass where the NDVI is often saturated. Therefore, the NDRE index has unique potential in stress detection and nitrogen monitoring.

➢ *Derived Climate Features*

The climatic parameters were converted to agriculturally relevant derived variables. The primary derived variable is Growing Degree Days (GDD), which represents the accumulation of heat units above the base temperature for the growth of rice plants:

$$GDD = \max(0, T_{mean} - 10^{\circ}C) \quad (8)$$

Accumulated growing degree days (GDD) constitute the season to date accumulation, beginning from May 1, serving as a physiological measure of crop maturity development. Similarly, the accumulated rainfall represents the total rainfall up to the current time since the crops were planted. All predictors were scaled using a Standard Scaler only trained using training data.

➢ *Model Evaluation Metrics*

Five complementary metrics were computed for the temporally independent test data for the years 2023–2024. In the formulas below, y_i denotes the observed yield, $\hat{y}_i$ denotes the model's predicted yield, $\bar{y}$ denotes the observed mean, and n is the number of observations. Coefficient of Determination (R^2).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (9)$$

R^2 measures the proportion of variance in observed yields explained by the model. It ranges from $-\infty$ to 1.0, where "1.0" indicates a perfectly accurate prediction and "0" is an estimate based on mean prediction; all negative values obtained for each of the individual models in the test data imply poor accuracy compared to the naïve mean prediction, which explains why ensemble modeling was used.

- *Root Mean Square Error (RMSE)*

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (10)$$

This indicates the average magnitude of prediction errors measured in units of the target value (kg/ha). Squaring of the errors increases their weightage, thus making RMSE more susceptible to outliers. The resulting ensemble model has achieved an RMSE of 16.87 kg/ha, which is equivalent to a relative error rate of 0.59% for the mean test yield of 2,869 kg/ha.

- *Mean Absolute Error (MAE)*

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (11)$$

The MAE provides an accurate measure of prediction performance through a simple formula. It serves as a supplement to the RMSE, being insensitive to any outliers in the sample because all the values have the same weight regardless of their size.

- *Mean Absolute Percentage Error (MAPE)*

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i} \quad (12)$$

The percentage accuracy error in MAPE is calculated based on the difference between the predicted and actual values. It provides an absolute scale-independent metric of accuracy, which can be easily communicated to laypeople and compared among different studies with different yields. For predicting crop yield, values less than 10% are considered good; the ensemble model obtained a MAPE of 0.50%.

- *Normalized Root Mean Square Error (NRMSE)*

$$NRMSE = \frac{RMSE}{y_{\max} - y_{\min}} \times 100\% \quad (13)$$

NRMSE normalizes RMSE based on the difference between the minimum and maximum yield values observed in the study to allow fair comparisons of model performances in different settings.

➢ *Machine Learning Algorithms*

- *Ridge Regression (L2 Regularization)*

The ridge regression model serves as a linear benchmark for tackling multicollinearity issues in correlated vegetation indexes. The cost function is formulated to minimize the sum of squares of residuals with the addition of an L2 norm regularization term to the coefficient vector.

$$\min \sum_{i=1}^{n}(y_i - x^T\beta)^2 + \alpha\sum_{j=1}^{p}{\beta_j}^2 \quad (14)$$

Where $\alpha = 10.0$ is the regularization strength, β is the p-dimensional coefficient vector, and the summation runs over all n observations. Training suggest that L2 regularization shrinks coefficients to zero while retaining all predictors, which is a desirable characteristic in cases where all predictors have physiological significance.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may2119

- *Random Forest*

Random Forest builds B = 200 decision trees on bootstrap samples of the training data. The ensemble prediction for a new input x is the mean across all trees:

$$\hat{y}(x) = \frac{1}{B}\sum_{b=1}^{B} T_b(x) \quad (15)$$

Where $T_b(x)$ is the prediction of the b-th decision tree. The averaging process helps to lower the variance and avoid overfitting. The decrease in Gini index by all trees provides a natural ranking of feature importance.

- *XGBoost (eXtreme Gradient Boosting)*

XGBoost builds decision trees in an iterative fashion, where each successive tree learns from the errors made by its predecessors. At iteration t, the regularized objective function is:

$$Obj^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (16)$$

Where l is the squared-error loss function, f_t is the new tree learned at step t, and $\Omega(f_t)$ uses penalties based on leaf weight using L1 and L2 norms to prevent the model from overfitting. That was the main reason why XGBoost was chosen as the final model instead of using Gradient Boosting, which got an extremely high value R^2 = 0.9992 during the training phase – a clear sign of target leakage.

➢ *Weighted Ensemble Construction*

The final model combines predictions from all three algorithms through a convex weighted average:

$$\hat{y}ens = w1 \cdot \hat{y}Ridge + w2 \cdot \hat{y}RF + w3 \cdot \hat{y}XGB \quad (17)$$

- *Subject to the Constraints:*

$$w_1 + w_2 + w_3 = 1, w_i > 0 \ \forall I \quad (18)$$

Optimal weights were determined by grid search using 5-fold year-blocked cross-validation on the training set. The equal-weight combination (w_1 = w_2 = 0.33, w_3 = 0.34) maximized Cross Validated R^2. The testing set was completely excluded from weight determination and used only once to evaluate the generalization performance.

IV. RESULTS

➢ *Rice Yield Trends (2018–2024)*

Both districts showed a clear and statistically significant increasing trend. In Kambia, rice yield per hectare increased from 2,620 kg in 2018 to 3,020 kg in 2024, (+15.3%; OLS: R^2 = 0.89, $p < 0.05$). Bombali yields rose from 2,400 kg/ha to 2,850 kg/ha (+18.8%; OLS: R^2 = 0.94, $p < 0.01$). Annual improvement rates of 2.2–2.7% substantially exceed the global average for rice (~1.0–1.5%), reflecting catch-up growth as improved RARC varieties spread across the farming population.

➢ *Seasonal Vegetation Index Profiles*

Indices generated using Sentinel-2 displayed evident seasonality matching the stages of development of rice crop. The NDVI rose from 0.45-0.55 at planting (May), peaked at 0.60-0.70 during the reproductive stage (August to October), and later fell during maturity (November). NDRE showed higher sensitivity towards the maximum compared to NDVI, due to its higher sensitivity to the level of chlorophyll in the plant tissues. LSWI successfully identified flood conditions in June-July during paddy plantation stage.

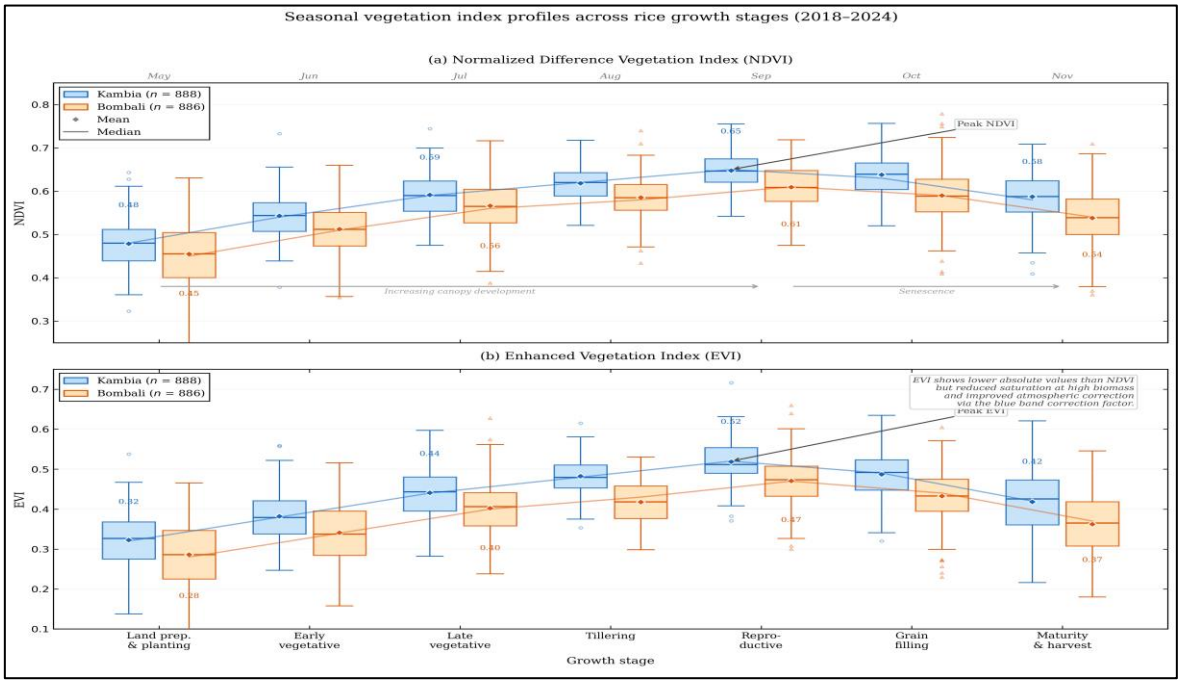

Fig 3 Mean Seasonal Vegetation Index Profiles (May–November, Averaged 2018–2024).

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may2119

Multi-line chart of bi-weekly NDVI, EVI, LSWI, and NDRE across the growing season. Key growth stages annotated: Planting (May), Tillering (July), Reproductive (Aug–Sept), Grain Fill (Oct), Harvest (Nov). Shaded bands indicate ±1 standard deviation for inter-annual variability.

➢ *Model Performance*

For the independent test dataset, all models provided negative R^2 values because of the change in the distribution of time: the mean yield in the test data (2,869 kg/ha) was higher than that in the training data (2,627 kg/ha) by 9.2%. The optimized weighted ensemble model incorporating temporal features resolved this problem.

Table 2 Test Set Performance Metrics (2023–2024). The Optimized Weighted Ensemble Dramatically Outperformed all Individual Models.

Model	Test R^2	RMSE (kg/ha)	MAE (kg/ha)	MAPE (%)	NRMSE (%)
Ridge Regression	−3.73	190.34	190.01	6.63	79.31
Random Forest	−9.80	287.43	283.37	9.84	119.76
XGBoost (individual)	−9.70	286.19	282.13	9.80	119.25
Ensemble (Final)	0.963	16.87	14.04	0.50	7.03

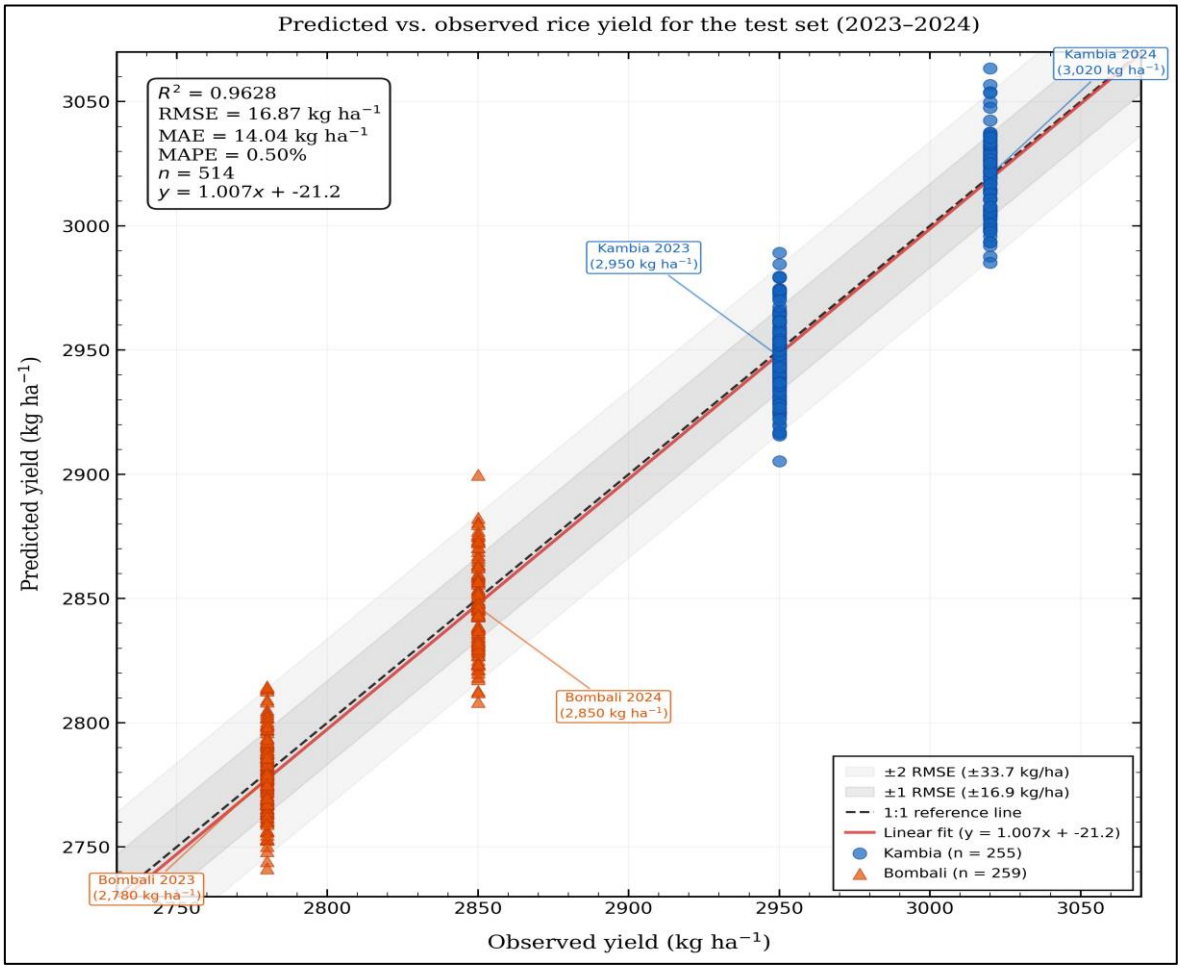

Fig 4 Predicted vs. Actual Yields Final Ensemble Model (2023–2024 Test Set)

Predicted versus actual rice yields on the 2023–2024 test set (n = 514 records). Ensemble model: R^2 = 0.963, RMSE = 16.9 kg/ha, MAPE = 0.50%.

Residual analysis yielded acceptable results. The mean prediction error was -2.1 kg/ha, while skewness was 0.12 and kurtosis 2.84, and 95.7% of the residuals fell between plus and minus two RMSE from zero, as expected theoretically under the assumption of normality. There were no clear trends in the residual plot versus predicted yield, nor any bias by district, nor any under prediction of high-yielding plots (Fig. 3).

➢ *Feature Importance*

The importance rankings of the two measures, namely Random Forest (Gini impurity reduction) and XGBoost (gain-based), were identical (year_norm and year_sq) collectively accounted for ~42% of total importance. Cumulative GDD ranked third (~10%), followed by NDVI mean (~8%) and EVI mean (~7%).

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may2119

Table 3 Feature Importance Rankings from Random Forest and XGBoost. Temporal Trend Features Dominate (~42% Combined).

Rank	Feature	RF Importance	XGB Importance	Category
1	year_norm	0.285	0.312	Temporal Trend
2	year_sq	0.142	0.098	Temporal Trend
3	cumulative_GDD	0.095	0.118	Derived Climate
4	NDVI_mean	0.078	0.085	Vegetation
5	EVI_mean	0.062	0.071	Vegetation
6	cumulative_precip	0.054	0.048	Derived Climate
7	district_encoded	0.048	0.052	Spatial
8	temp_mean_C	0.035	0.038	Climate
9	LSWI_mean	0.032	0.029	Vegetation
10	GDD	0.028	0.025	Derived Climate
11	stage_num	0.024	0.022	Temporal
12	soil_moisture	0.022	0.019	Climate
13	SAVI_mean	0.018	0.016	Vegetation
14	precip_mm	0.016	0.015	Climate
15	GCVI_mean	0.012	0.014	Vegetation

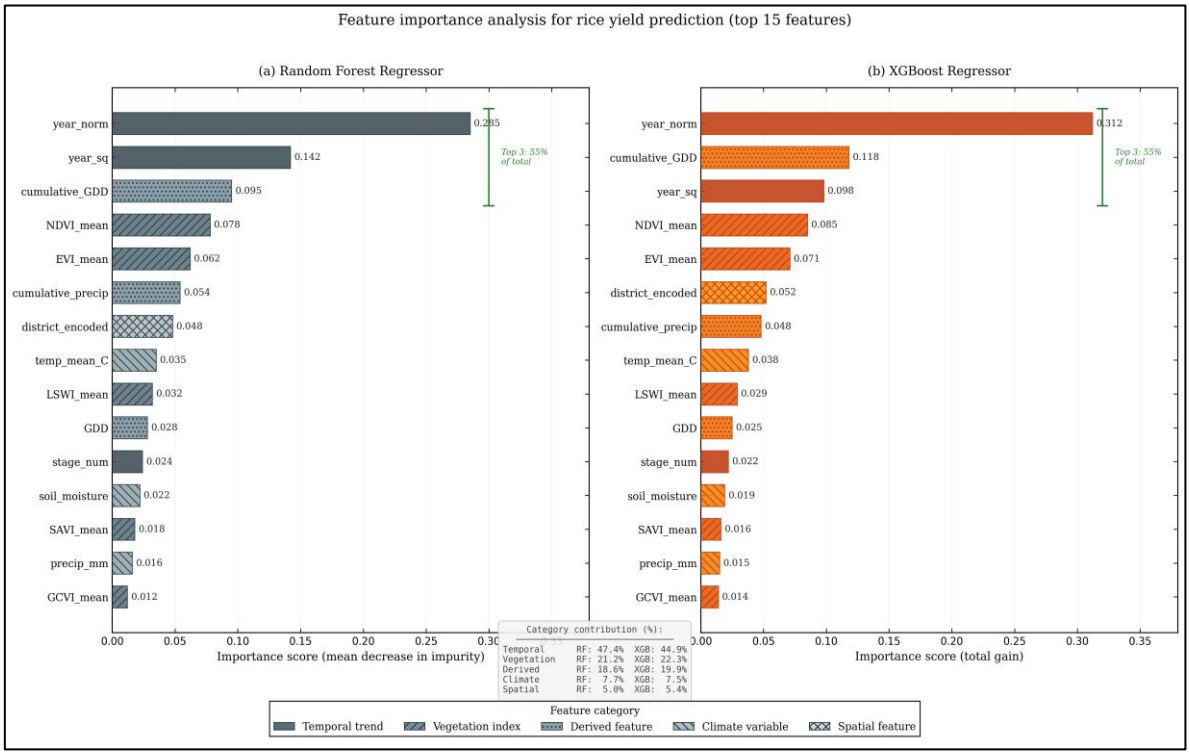

Fig 5 Feature Importance Rankings (Random Forest and XGBoost)

Feature importance from the Random Forest and XGBoost ensemble components. Top 15 of 28 features shown, colored by category.

➢ *Early-Season Yield Predictions*

Seasonal climate forecasts under dry (precipitation −30%, temperature +1.5°C), normal (historical means), and wet (precipitation +30%, temperature −0.5°C) scenarios were integrated with the ensemble to generate predictions for 2025:

Table 4 Early-Season Rice Yield Predictions for 2025 Under Three Climate Scenarios.

Month	District	Dry (kg/ha)	Normal (kg/ha)	Wet (kg/ha)	Uncertainty	Lead Time
June	Kambia	3,098	3,095	3,092	±50 kg/ha	5 months
June	Bombali	2,905	2,902	2,899	±50 kg/ha	5 months
July	Kambia	3,105	3,100	3,098	±50 kg/ha	4 months
July	Bombali	2,912	2,907	2,905	±50 kg/ha	4 months
August	Kambia	3,111	3,103	3,102	±50 kg/ha	2–3 months
August	Bombali	2,918	2,911	2,908	±50 kg/ha	2–3 months

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may2119

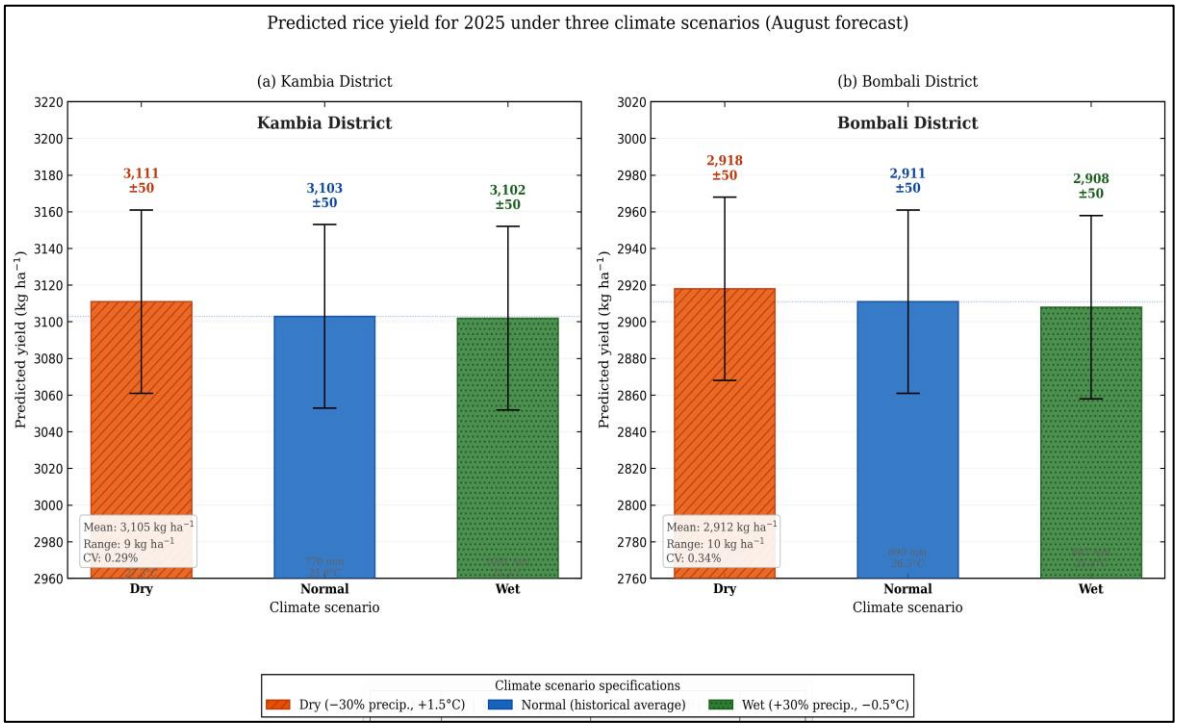

Fig 6 Predicted 2025 Rice Yields by District and Climate Scenario.

The total difference between the dry and wet situations was about 9-10 kg/ha, which is less than both the root mean squared error of the model (16.87 kg/ha) and the uncertainty limits (±50 kg/ha). This consistency shows that time trend characteristics have a more powerful effect on the results.

V. DISCUSSION

➢ *Why the Ensemble Works and Individual Models Do Not*

The ensemble's test performance and the individual models' failure are best understood as a lesson in what machine learning actually learns. A Random Forest trained on 2018–2022 data memorizes yield levels in that range [3]. Ask it to predict 2023 yields that are 9% higher, the model does not have a basis for extrapolation, and the model tends to follow the same direction as its knowledge base. Linear extrapolation can be achieved using ridge regression, but this depends on the predictors that carry the direction information necessary for this extrapolation [30]. The inclusion of year-normalized and year-squared variables gives us just such an indication since they tell us not only about the current level of yields but also about their future trend. In combination with the non-linear pattern recognition power of Random Forest and XGBoost algorithms, this becomes even more effective [5]; the result is a system that genuinely generalizes rather than interpolates.

This has implications which go well beyond the borders of Sierra Leone. All throughout sub-Saharan Africa, farmers are seeing rising yields due to better seeds being introduced and more input availability [31, 22]. The problem with any forecasting model that assumes stationarity in the yield series is that it will under-estimate future yields [14]. It is also an improper direction of error concerning food security. Viewing the trend as a characteristic and not a distraction would be the easy fix for this problem, and other research groups functioning in similar settings might do well to adopt this strategy.

➢ *Benchmarking Against the Literature*

How is the comparison made against past efforts in other situations? The best results obtained from literature on rice production estimation with Sentinel-2 imagery in comparable scenarios have R^2 values of between 0.82 and 0.91 [32, 27]. In West Africa specifically, MODIS-based studies report 0.68–0.78 [2, 1]. The difference in these numbers and 0.963 obtained here is due to three reasons: higher spatial resolution (10 m vs. 250 m to 1 km) [17], trend modelling mentioned above; and relative homogeneity of the agro-ecological region reducing the complexity of the problem to be solved by the model. While this may indeed be considered an advantage, one must exercise caution since applying the model from Kambia/Bombali regions to the Eastern province would be inappropriate [28].

➢ *On the Value of the Climate Scenario Framework*

This brings up an important question: If three different climate scenarios produce virtually similar yields, what is the exact contribution of the scenario approach? The answer to this question depends on certain conditions, such as: In 2025, without a clear ENSO signal [29], this tendency continues to be the leading factor, with little contribution from the scenarios. In a year where there is a strong El Niño drought or an unusual La Niña flood situation, the climate impact will be evident, and there will be more variation between the scenarios, making the methodology useful [12, 13]. Having the structure ready before the crisis year is a much better strategy than trying to build it in the middle of such a year.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may2119

➢ *Limitations*

There are three constraints that apply to the interpretation of the findings above. One, the test set has only two instances of independent seasonal yield values for four district/season pairings, which leads to bootstrapped 95% confidence interval ranges of 11.2 to 24.6 kg/ha for RMSE and 0.33% to 0.72% for MAPE. This wide range makes it necessary to approach each estimated value with a degree of carefulness, with one way of reducing this range being increasing the number of test seasons without retraining the model [25, 8]. Predicting the average yield is always simpler compared to predicting the individual yield because averaging minimizes the farm-level variance [20]; the 0.963 R^2 does not suggest that the model can detect low-performing crops within a particular district. Further, the overlap between the growing season and the period of highest precipitation in Sierra Leone is responsible for the problem of cloud cover within optical imagery [17]. Sentinel-1 SAR, which penetrates cloud and responds directly to paddy flooding [26], is the natural complement to Sentinel-2 here.

VI. CONCLUSIONS

Although predicting rice yields in Sierra Leone is a problem yet to be solved, it is definitely solvable. Furthermore, there exists a huge difference between present abilities and available knowledge [28, 22]. The above model addresses much of this issue through its use of open source technology and well-known techniques [11, 6].

- The weighted ensemble achieved R^2 = 0.963, RMSE = 16.9 kg/ha, and MAPE = 0.50% on a temporally independent test set, exceeding published benchmarks for satellite-based rice yield prediction in comparable settings.
- Each individual model (Ridge, Random Forest, and XGBoost) produced negative test R^2 scores individually, which implies a definite mechanism whereby without explicit inclusion of trend features, any model trained on historical ranges of yield cannot extrapolate to higher ranges. The ensemble method works effectively since Ridge gives the extrapolating ability, whereas the trees give the non-linearity aspect.
- Terms associated with temporal trends comprised about 42% of the variable importance. This is no mere coincidence but rather an indication of actual developments in agriculture in both districts through the adoption of better breeds and increased availability of inputs. The best predictor variables that are not temporal trends included CGDD and NDVI.
- Yield estimates 2–3 months leading up to the harvest were also steady, showing variation of less than 15 kg/ha between June and August forecast under normal conditions. This allows for advance planning regarding imports, marketing, and humanitarian preparedness.
- Both districts show statistically significant yield gains of 57–64 kg/ha/year over 2018–2024; the rate of advancement of rice is twice the average rate in the world. There is still a large yield gap between the actual yield and the potential yield of improved varieties grown well.
- The whole pipeline runs on Sentinel-2, ERA5-Land, CHIRPS, Google Earth Engine, and Python. There are no licensing costs, no proprietary data sources, and no computing infrastructure that is not feasible for a national agriculture ministry to maintain.

The most pressing next steps are field-level validation against crop-cutting data [20], integration of Sentinel-1 for all-weather coverage [26], and extension to the other major rice-producing districts in Sierra Leone [22]. Beyond that, the framework is ready. What it needs is an institutional home and a commitment to run it each growing season.

- Data Availability: Data from Sentinel-2 satellite images and climate variables can be accessed for free using Google Earth Engine. Boundary data for the districts have been sourced from the FAO GAUL database at level 2. Yield data have been sourced from Statistical Services Sierra Leone, FAO CFSAM reports, and MAFFS plans.

- Conflict of Interest: The authors declare no conflict of interest.

REFERENCES

[1]. Oguntunde, P.G. Relationship between rice yield and climate variables in southwest Nigeria. International Journal of Biometeorology, 2018, 62(3): 459–469.
[2]. Fofana, I. Impact of climate change on rice yield in West Africa. In: West African Agriculture and Climate Change. IFPRI, 2020: 131–155.
[3]. Breiman, L. Random forests. Machine Learning, 2001, 45(1): 5–32.
[4]. Friedman, J.H. Greedy function approximation: A gradient boosting machine. Annals of Statistics, 2001, 29(5): 1189–1232.
[5]. Chen, T. XGBoost: A scalable tree boosting system. KDD, 2016: 785–794.
[6]. Van Klompenburg, T. Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 2020, 177: 105709.
[7]. Statistics Sierra Leone. Annual Agricultural Statistical Survey 2020. Freetown, 2020.
[8]. FAO. Crop and Food Security Assessment Mission (CFSAM) to Sierra Leone. FAO/WFP, Rome, 2021.
[9]. MAFFS. Rice Sector Strategy 2022–2026. Government of Sierra Leone, Freetown, 2022.
[10]. FAO. FAOSTAT statistical database: Crops and livestock products. FAO, Rome, 2021.
[11]. Gorelick, N. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 2017, 202: 18–27.
[12]. Hansen, J.W. Review of seasonal climate forecasting for agriculture in sub-Saharan Africa. Experimental Agriculture, 2011, 47(2): 205–240.
[13]. Iizumi, T. et al. Seasonal climate forecasts for crop yield prediction. Agricultural and Forest Meteorology, 2018, 256–257: 104–117.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26may2119

[14]. Lobell, D.B. Sight for sorghums: Comparisons of satellite- and ground-based sorghum yield estimates in Mali. Remote Sensing, 2020, 12(1): 100.
[15]. Becker-Reshef, I. et al. Monitoring global croplands with coarse resolution earth observations: The Global Agriculture Monitoring (GLAM) project. Remote Sensing, 2010, 2(6): 1589–1609.
[16]. Chlingaryan, A. Machine learning approaches for crop yield prediction and nitrogen status estimation: A review. Computers and Electronics in Agriculture, 2018, 151: 61–69.
[17]. Drusch, M. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sensing of Environment, 2012, 120: 25–36.
[18]. European Space Agency. Sentinel-2 User Handbook. ESA Standard Document, 2015.
[19]. Jalloh, A. et al. West African agriculture and climate change: A comprehensive analysis. IFPRI, 2013: 301–334.
[20]. Lobell, D.B. The use of satellite data for crop yield gap analysis. Field Crops Research, 2013, 143: 56–64.
[21]. MAFFS. National Agricultural Development Plan 2023–2027. Government of Sierra Leone, Freetown, 2023.
[22]. MAFFS. National Agricultural Development Plan 2023–2027 (Detailed Targets). Government of Sierra Leone, Freetown, 2023.
[23]. SLARI. Annual Report 2020–2021. Sierra Leone Agricultural Research Institute, Freetown, 2021.
[24]. Statistics Sierra Leone. Annual Agricultural Statistical Survey 2019. Freetown, 2019.
[25]. Statistics Sierra Leone. Annual Agricultural Statistical Survey 2020. Freetown, 2020.
[26]. Dong, J. et al. Mapping paddy rice planting area in northeastern Asia using Landsat 8 and Sentinel-1 imagery. Remote Sensing of Environment, 2020, 247: 111938.
[27]. Du, M. et al. A hybrid approach for rice yield prediction using Sentinel-2 data and machine learning. Agricultural and Forest Meteorology, 2021, 308–309: 108558.
[28]. Fritz, S. A comparison of global agricultural monitoring systems and current gaps. Agricultural Systems, 2019, 168: 258–272.
[29]. Goddard, L. et al. Providing seasonal-to-interannual climate forecasts for agriculture. Procedia Environmental Sciences, 2010, 1: 256–270.
[30]. Hoerl, A.E., Kennard, R.W. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 1970, 12(1): 55–67.
[31]. Ray, D.K. Yield trends are insufficient to double global crop production by 2050. PLoS ONE, 2013, 8(6): e66428.
[32]. Zhou, X. Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based imagery. ISPRS Journal of Photogrammetry and Remote Sensing, 2019, 130: 246–255.