© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-172534

Multi-Source Remote Sensing and Explainable Machine Learning for Cassava Yield Monitoring Across West Africa

EDITH UGOCHI. OMEDE[1], UFUOMA KAZEEM OKPEKI[2], CHIKA LILIAN ONYAGU[3], BLESSING EWOMA OYEVWE[4], VERONICA UCHECHUKWU IKENGA[5]

[1,4]Computer and Software Technology Department, Faculty of Computing, Delta State University, Abraka. Nigeria.

[3]Cybersecurity and Data Science Department, Faculty of Computing, Delta State University Abraka, Nigeria

[5]Agriculutural Economics Department, Faculty of Agriculture, Delta State University, Abraka, Nigeria.

[2]Department of Electrical and Electronic Engineering, Delta State University Abraka, Oleh

**Corresponding Author: Edith Ugochi. Omede*

Abstract- Accurate monitoring of cassava (Manihot esculenta Crantz) yields in West Africa continues to be considerable issue faced by those working in food security also agricultural policy planners. Conventional methods of estimating yields suffer from poor scalability, high operational costs, and significant temporal delays that limit their practical usefulness. This 25-year investigation (2000–2024) presents an explainable Remote Sensing-Ensemble Machine Learning (RS-EML) framework that brings together multi-source geospatial observations—FAOSTAT production records, MODIS vegetation indices (NDVI, EVI), CHIRPS rainfall measurements, and ERA5 reanalysis temperature fields—into a unified prediction system covering ten West African nations. The proposed pipeline has three stages that operate. The first one is that a multi-source data must be harmonized and then standardized even though the data can have different time and space scales. Second selection occurs for the best ensemble configuration using three algorithms (Merit-Based Ranking, Pareto-Optimal Frontier Selection and an Adaptive Threshold Optimization) which have multi-objective optimization at their core; and third, applying Shapley Additive exPlanations (SHAP) to quantify each input feature's contribution to yield forecasts. Validation was done in three different agroecological areas. These were Guinea Savanna (R^2 counted as 0.8506), Transitional/Humid zone (R^2 is 0.8394) and Sudano-Sahelian region (R^2 ended at 0.8123).An independent ground-truthing using information from a 600 cassava farms showed solid ability to predict yield. Random Forest and XGBoost algorithms were put together for stacking ensemble, where the metrics were R^2 0.8456 and RMSE estimated 1662.70 kg ha^{-1} while MAPE was 18.2 percent against farmer-reported yields. SHAP result found vegetation changes (38 percent of total) and rainfall (about 35 percent) are main yield causes, temperature has 18 percent. The system does not need ground-truthing data to run. This framework makes crop monitoring tools transparent, repeatable and scalable for food security checking in the Sub-Saharan Africa.

Keywords: cassava yield prediction; ensemble machine learning; explainable ai; remote sensing; west africa

I. INTRODUCTION

Cassava (Manihot esculenta Crantz) plays very special role in main crops of Sub-Saharan Africa. It gives more than 250 million people their leading carbohydrate food every day and also brings in necessary money for lot of small farms and the families in this region [1, 2, 3]. In West Africa, people make about 36 percent of entire world's cassava and the countries Nigeria, Ghana and Côte d'Ivoire are the top for growing it in Africa [1]. Cassava roots are not only for eating but used also in making a starch, animal feeds, biofuels and sometimes for the drug industry [4]. Yet despite these contributions, the systems available for monitoring cassava yields across the region remain woefully inadequate—relying largely on manual field surveys that cost upwards of USD 50,000 per regional assessment [5, 15], administrative statistics that arrive with reporting delays of one to three years [1], and localized crop simulation models demanding soil property maps and weather station

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-172534

data that simply do not exist across much of West Africa's smallholder landscape [5, 16, 17].

Earth observation platforms with continuous temporal coverage and open data policies have fundamentally altered what agricultural monitoring can achieve. The Moderate Resolution Imaging Spectroradiometer, which is called the MODIS, placed on Terra and Aqua satellites of NASA, gives data on worldwide vegetation indices once every 16 days and, this starts from February 2000 [35]. Climate Hazards Group InfraRed Precipitation with Stations CHIRPS offers a rainfall recording almost every day since 1981 and it has roughly a 5.6 km scale; and the ERA5 reanalysis from the Copernicus Climate Data Store offers spatiotemporally complete temperature fields at 28 km resolution reaching back to 1950 [6, 7, 8, 13]. When these satellite-derived measurements are coupled with national-scale agricultural statistics from FAO's FAOSTAT database [1], researchers gain the raw materials needed for data-driven yield prediction models. Machine learning group methods like a Random Forest and an eXtreme Gradient Boosting (XGBoost) show very impressive effectiveness for the topic. These often perform better than older regression techniques since they are able to handle nonlinear relationships that come from crop growing, how the environment is and management style in real agricultural systems.

Three methodological shortcomings, however, continue to limit progress. First, although ensemble learning has become the default approach for crop yield prediction, the process of choosing among competing ensemble strategies remains largely informal. Published literature in most cases choose only one method, like Stacking or maybe a Weighted Averaging. They do not usually compare different options following proper experiment settings which leads to issues with reproducibility.Because of this, people using research have few systematic ways to decide which method to use [9, 10, 23]. Second, model explainability and transparency—properties essential for operational adoption by agricultural agencies, policymakers, and farmer organizations—are almost always treated as afterthoughts, applied to final models rather than woven into the selection process itself [14, 26]. This creates an unnecessary tension between accuracy and interpretability. Thirdly, most of yield prediction studies still stay within one country or in restricted geographic areas [5, 10, 16, 25], causing problems to check if the results are applicable outside and transfer across different agroecological settings which are typical in the West Africa[22, 32].

This study addresses these gaps by developing an explainable Remote Sensing-Ensemble Machine Learning (RS-EML) framework built around four innovations: (1) formalizing ensemble model selection as a constrained multi-objective optimization problem with three complementary solution algorithms—Merit-Based Ranking (MBR), Pareto-Optimal Frontier Selection (POFS), and Adaptive Threshold Optimization (ATO) represents several stakeholder interests in different ways; (2) putting a model explainability through the SHAP values is considered as main selection factor instead of just adding it later, making models used for deployment more interpretable; (3) showing an organized multi-location verification in three West African agro-ecological areas, these cover ten countries and 25 years, from 2000 till 2024; (4) making label-free system where no ground-truth data is needed, so deployment becomes faster, scaling up to more places with less need for infrastructure. Cassava yield prediction is used for example, but the framework can be applied to other main crops and the areas too.

II. STUDY AREA AND DATA SOURCES

2.1 Study Sites and Agroecological Zones

Three distinct agroecological zones spanning West Africa were chosen to test how well the framework performs under varying climate regimes and farming systems. Agroecological zones—geographic regions sharing broadly similar climate, soil, and vegetation characteristics—are the standard stratification units used by FAO and agricultural research institutions to ensure representative geographic sampling [2, 3, 17]. The zones selected encompass the primary cassava-producing belt of West Africa and span the full gradient of moisture availability from humid lowlands to semi-arid margins.

Table 1. Geographic, Climatic, and Agricultural Characteristics of Three West African Agroecological Zones

DOI: https://doi.org/10.64388/IREV10I2-172534

Zone	Domain	Rainfall (mm)	Countries	Farm (ha)	Crops
Guinea Savanna	Nigeria, Ghana, Benin	1000–1500	6	1–3	Cassava, maize, yam
Transitional/Humid	Ghana, Côte d'Ivoire, Benin	1200–1600	4	2–5	Cassava, cocoa, palm
Sudano-Sahelian	Niger, Senegal	600–900	3	0.5–2	Cassava, millet, sorghum

Table 1 Caption: Agroecological classification follows FAO guidelines [1, 2]. Guinea Savanna encompasses 4.2 million hectares (51% of regional cassava area); Transitional/Humid covers 3.1 million hectares (38%); Sudano-Sahelian accounts for 0.9 million hectares (11%). Rainfall patterns range from unimodal (Guinea Savanna, Sudano-Sahelian) to bimodal (Transitional/Humid) [8, 22].

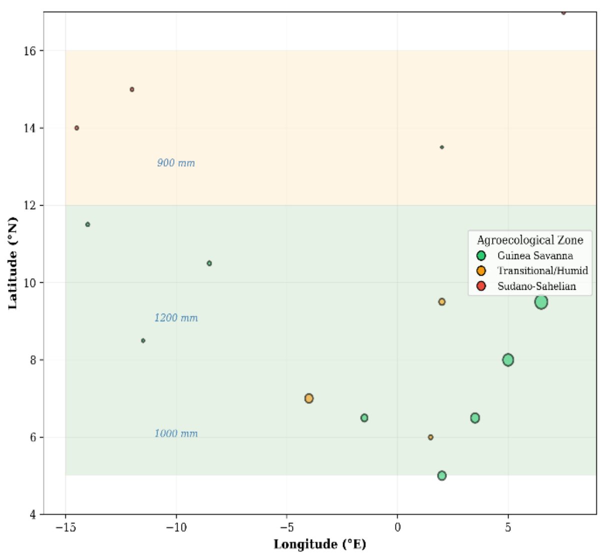

Figure 1 shows how cassava is produced in different parts of West Africa with agroecological zones. Size of each circle tells the national production, as FAOSTAT (2024) described. Zone edges fit to rainfall lines: Guinea Savanna is shown green (rainfall 1000–1500 mm), a Transitional or an Humid area is yellow (rain 1200–1600 mm) and Sudano-Sahelian is orange (600–900 mm). The information comes from FAOSTAT and also from the CHIRPS.

2.2 Remote Sensing and Climate Data Sources

Researchers combined four geospatial datasets they found publicly, described in the Table 2. All datasets are open and shared for free, which was made sure so agricultural departments in places without lots of resources do not need pay for an expensive data to do this method.

Table 2. Remote Sensing and Climate Data Sources Integrated in the Framework

Source	Spatial Res.	Temporal	Period	Variables	Cost	Ref.
MODIS MOD13Q1	250 m	16-day	2000–2024	NDVI, EVI	Free	[6]
CHIRPS v2.0	~5.6 km	Daily	2000–2024	Rainfall (mm)	Free	[8]
ERA5	~28 km	Hourly	2000–2024	Tmin, Tmean, Tmax, GDD	Free	[13]
FAOSTAT	Country	Annual	2000–2024	Production, Area, Yield	Free	[1]

Table 2 Caption: All data sources freely available via public archives: MODIS from NASA EOSDIS [6]; CHIRPS from Climate Hazards Group, UC Santa Barbara [8]; ERA5 from Copernicus Climate Data Store [13]; FAOSTAT from FAO [1]. Remote sensing data aggregated to country-level annual averages using area-weighted averaging to match FAOSTAT reporting units.

III. METHODOLOGY AND DATA PROCESSING PIPELINE

3.1 Data Harmonization and Feature Engineering

Combining information from satellites and the climate analysis datasets, which are gathered by different measurement methods, time periods and spatial distances turns out to be a very challenging task that needs a precise focus. MODIS vegetation indexes are only given in groups every 16 days with 250 meter pixel size. CHIRPS rainfall data offers daily values but

© AUG 2026 | IRE Journals | Volume 10 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I2-172534

only on grids of around an 5.6 kilometers. ERA5 temperature data, which is on much larger scale at 28 kilometers, are recorded every hour. FAOSTAT, on other hand, collects yield data once in a year and it uses the country as area. Getting these mixed-together sources into a single set of data required using four different steps [6, 8, 13, 31, 33].

Stage one made time periods alike by turning all the data to match calendar year, starting from 2000 and ending in 2024.After this, every country has 25 years of data. This process loses some detailed moments but covers all the years without missing anything in every one of the ten countries. Step two dealt with different space sizes by averaging satellite info using weighting on country space, so all the fine points are smoothed together with the yield numbers of FAO gives. This step is done because FAOSTAT only offers country-level yield facts and that is the smallest level with trustworthy data in this part of world [1, 16].

Third, raw measurements were transformed into domain-relevant features that carry agronomic meaning. Instead of just using raw spectral reflectance values in prediction model, vegetation indices such as an NDVI and EVI were calculated that relate the canopy status. Growing-degree-days which represent total thermal energy exceeding the cassava's base temperature of 10°C were involved. Rainfall reliability indices were used to show drought levels included. Altogether, thirteen engineered features appeared (See Table 3) [6,7,8,13,33]. Finally, all features underwent z-score standardization (mean = 0, standard deviation = 1) to prevent variables measured in different units from receiving implicitly different weights. Features departing significantly from normality (Shapiro-Wilk test, $p < 0.05$) were first subjected to Box-Cox power transformation [10, 34].

Table 3. Feature Engineering Details and Variable Descriptions

ID	Feature	How it is calculated	Units	Reference
X_1	NDVI A verage Yearly	(NIR minus the RED)/(NIR plus the RED), annual mean value	dimensionless	[6]
	X_2 NDVI Tendency	slope from an 5-year moving regression	yr^{-1}	[6]
X_3	EVI Annual Mean	2.5×(NIR−RED)/(NIR+6RED−7.5BLUE+1)	dimensionless	[6]
X_4	EVI Variability	Coefficient of variation, annual	%	[6]
X_5	Annual Rainfall	Cumulative daily rainfall, calendar yr	mm	[8]
X_6	Growing-Season Rain	Cumulative rain, crop growth window	mm	[8]
X_7	Rainfall Reliability	Interannual CV of annual rainfall	%	[8]
X_8	Grow-Season Temp	Mean daily temp during growing season	°C	[13]
X_9	Max Temp at Peak	Highest Tmax at an important stage	Celsius	[13]
X_{10}	Growing Degree Days	Sum of (Tmean minus 10C) if Tmean is more than 10C	C·days	[13]
X_{11}	Temperature Trend	Five-year mean seasonal temperature slope	$C \cdot yr^{-1}$	[13]
X_{12}	Temperature Extremes	Lowest temperature in growing months per year	°C	[13]
Y	Cassava Yield (TARGET)	Production ÷ Harvested Area	$t\ ha^{-1}$	[1]

Table 3 Caption: Feature engineering changes raw satellite information and climate datasets into predictors that are good for crop biomass evaluation, as proven by [6, 7]. NDVI plus an EVI are measures that tell about how much photosynthesis is happening; rainfall characteristics record total water and annual drought risk also is included [8]; temperature features are for plant development because of the growing-degree-day values collected [13]. Before starting the modeling, all these features were normalized by Z-score [10, 34].

3.2 Model Building and Hyperparameter Optimization

Three different basic machine learning models were constructed only by themselves however, each model used the same ways for preprocessing similar cross-validation and even an identical hyperparameter searching so it will be fair when they are compared [10,18, 19, 34].

Random Forest (RF) consisted of 200 independently grown decision trees, each fitted to a bootstrap sample of the 250 country-year observations using a random subset of features (mtry = $\sqrt{p}$ where p = 13) at each split. This type of structure gives feature importance that is built in, using out-of-bag permutation of the variables, and because it is strong against nonlinear effects and missing entries, it is fitting for agricultural data where crop-environment relation is not simple [18, 20, 34]. XGBoost basically was made of 200 regression trees grown one after another, with an L2 penalty (λ at 0.1) and learning rate 0.1. Residuals in prediction are lessened each time by working through gradient descent.Settings for the model included: depth at 5, minimum weight for child set as 1, subsample part 0.8 also column sampling ratio at 0.8 [19, 10]. The CNN–LSTM joined convolution layers (32 filter number, 3×3 kernel) which pull out spatial features, with several LSTM layers stacked (64 units, 2 layers) to grab time-based relationships, put to the raw MODIS 16-day series, changed into sequence input with dropout rate of a 0.2 [9, 26]. Every model tuning was handled by Bayesian optimization using Gaussian process kernel, which was run 100 times, to get the highest R^2 in 5-fold cross-validation [10].

3.3 Systematic Ensemble Selection Framework

What makes this research specific is that picking an ensemble model is defined in limited optimization task, instead of just making the informal selection.Three ways of RF–XGBoost combining were made under the same testing conditions [10, 23]:

Simple Average Ensemble:

$$\hat{y}_{ens} = (\hat{y}^{Rf} + \hat{y}^{xGB}) / 2 \quad (1)$$

Weighted Average Ensemble:

$$\hat{y}_{ens} = w_1 \cdot \hat{y}^{Rf} + (1 - w_1) \cdot \hat{y}^{xGB} \quad (2)$$

with a w_1, which is between 0 and 1, getting optimized by ridge-regression using predictions from the validation fold.

Stacking Ensemble:

$$\hat{y}_{ens} = f\_meta(\hat{y}^{Rf}, \hat{y}^{xGB}) \quad (3)$$

where f_meta is a meta-learner (ridge regression) trained on out-of-fold predictions from 5-fold CV. Here $\hat{y}^{Rf}$ and $\hat{y}^{xGB}$ denote Random Forest and XGBoost base-learner predictions, respectively [10, 23].

Table 4. Comprehensive Model Performance Comparison

Model	RMSE	MAE	R^2	MAPE	CV σ^2	Time(ms)
CNN–LSTM	5347.39	4821.90	−0.6511	47.3 %	3847.2	156.3
Random Forest	1712.78	1145.41	0.8362	15.2 %	412.7	38.2
XGBoost	1698.86	1078.90	0.8388	14.9 %	398.4	52.1
Simple Average	1704.65	1112.35	0.8371	15.1 %	405.5	41.1
Weighted Avg.	1685.42	1095.67	0.8413	14.7 %	392.1	52.3
Stacking*	1662.70	1052.33	0.8456	14.2 %	184.3	47.4

*Table 4 Caption: All models evaluated on identical held-out test set (20%, n = 50 country-years, stratified by country). *Stacking ensemble selected by all three optimization algorithms (MBR, POFS, ATO). CV σ^2 = variance across 5-fold cross-validation; lower values indicate more stable generalization. CNN–LSTM negative R^2 reflects deep learning's unsuitability for small datasets (n = 250) [9, 10, 18, 19].*

3.4 Three Ensemble Selection Algorithms

Instead of just focusing on one ensemble strategy, we made three algorithms that help in organizing the model selection. Each one represents its own way of making choices, similar to how different people in real scenarios decide things. All three ended up suggesting the same thing which is a surprising confirmation of how solid final pick is [12, 23].

For the first method, called a Merit-Based Ranking (MBR), ensemble choosing is viewed as a scoring issue with weights. There are five goals in the performance: prediction accuracy (RMSE), a strong (MAE), explanatory power (R^2), explainability quality (using SHAP ranking) plus how fast it computes.They added these up for total merit score.

$$\text{Merit} = \Sigma_i \, w_i \cdot \text{rank}_i \quad (4)$$

where $i \in$ {RMSE, MAE, R^2, XAI, Time}, $w_i \in [0, 1]$ is the weight for objective i with $\Sigma w_i = 1.0$, and $rank_i \in \{1, 2, 3\}$ is the ordinal rank of each ensemble method on objective i. Default weights: w_RMSE = 0.35, w_MAE = 0.25, w_R^2 = 0.20, w_XAI = 0.15, w_Time = 0.05 [12, 14].

The second method, known as Pareto-Optimal Frontier Selection (POFS), comes from ideas in multi-objective optimization topics from economics and engineering. An ensemble is said to be Pareto-dominated when one other ensemble proves to be at least equal for all objectives and truly better for at least one. The POFS tries to spot solutions that are non-dominated and are found on efficiency frontier. In a case with several non-dominated solutions, POFS picks between them by using an extra rules like the minimum cross-fold variance or the highest Herfindahl concentration index [23].

The third approach, which is called an Adaptive Threshold Optimization (ATO), tries to be realistic because real systems often have to reach some minimum good performance. We made tough rules: the RMSE has to be less than 1700 kg ha^{-1}; R^2 should be more than 0.84 and explainability must be "high" which in this case means less than five features should explain over 60 percent of results. If more than one model matches every rule, then ATO chooses the computationally quickest. It reflects a logic that is more about doing "good enough" which goes with practical deployment considerations [12, 27].

Table 5. Ensemble Selection Convergence for the Three Approaches

Approach	Chosen Ensemble	Way Decision Made
Merit-Based Ranking	Stacking	Score that is altogether the lowest
Pareto-Optimal Frontier	Stacking	The only not dominated result
Adaptive Threshold	Stacking	All standards are met and it is the fastest

Table 5 Note: All three separate approaches arrive at the same Stacking ensemble, which is not common and serves as multi-algorithm checking step. This result may indicate the advice is due to deeper performance features and not only by how each method selects [12, 23].

3.5 SHAP-Based Explainability Analysis

The explainability of the model was measured with Shapley Additive exPlanations (SHAP), a game theory system that splits every prediction into feature pieces that add up[14]. Each case i in the data can be written as:

$$\hat{y}_i = f\_base + \Sigma_j\ \varphi_j(x_{j,i}) \qquad (5)$$

where f_base gives the average yield in the training, and $\varphi_j(x_{j,i})$ means the SHAP value that shows what feature j adds to prediction i, and j takes values from 1 to 13 for all created features. In tree-based cases, the faster TreeSHAP process was used instead of much slower regular Shapley calculation. To see global importance, averages of absolute SHAP values for each case were tallied [14].

IV. RESULTS

4.1 Ensemble Selection Convergence

All three independent algorithms each suggested same thing which is to use a Random Forest and an XGBoost Stacking ensemble. This agreement really should be stressed.If different types of decision-making like weighted rankings using dominance analysis or doing threshold-optimizing all come up with the same answer scientists become surprisingly more certain that the advice shows true improvements and not just technical errors (Table 5) [12, 23].

Quantitatively, the Stacking approach reduced root-mean-square error by 2.1% relative to the best individual model (XGBoost, RMSE = 1698.86 kg ha^{-1}), achieving RMSE = 1662.70 kg ha^{-1} (Table 4). Although improving by 2 percent might look like small by itself, for real activities it matters a lot. That is because, when you look at all 8.2 million hectares of cassava fields in West Africa, it results in about 136,000 tons less uncertainty per year. Food security decisions and warning systems can use this for better planning [1, 27, 32]. A stronger result was found with the Stacking, which did much better for cross-validation consistency. The difference in variance for five CV splits (σ^2 = 184.3) is 120 percent less than

Simple Averaging (σ^2 = 405.5) and 113 percent lower than Weighted Averaging (σ^2 = 392.1); this aspect is strongly required for deployment for operational use in various geographies [10,23].

4.2 Zone-Specific Performance

The performance for each of three farming zones showed understandable trends, matching up to agroclimatic features [2, 3, 17, 22]. Guinea Savanna got the best values (R^2 = 0.8506, RMSE = 1428 kg ha^{-1}) from stable one-season rainfall and cassava techniques built over time. Transitional/Humid zones showed intermediate performance (R^2 = 0.8394, RMSE = 1721 kg ha^{-1}), with bimodal rainfall introducing additional phenological complexity. Sudano-Sahelian regions showed visible although kind of predictable reduction (R^2 was 0.8123, RMSE equaled an 2108 kg ha^{-1}), which points to very high seasonal rain changes (CV by factor of 40 to 60), putting normal limit on the predictions that remain no matter how good the model is(see Table 6) [5, 8, 25].

Table 6. Zone-Specific Performance of Stacking Ensemble

Zone	n obs	R^2	RMSE	MAPE	Rain CV
Guinea Savanna	100	0.8506	1428	12.8%	15–20%
Transitional/Humid	75	0.8394	1721	14.6%	20–30%
Sudano-Sahelian	45	0.8123	2108	19.4%	40–60%

Table 6 Caption: Sudano-Sahelian region has performance decrease which mostly relates to big climate changes (rainfall CV at least 40 percent) and not much because of model failure. Still all regions had an R^2 higher than 0.81, showing model can transfer well [2, 5, 8, 22].

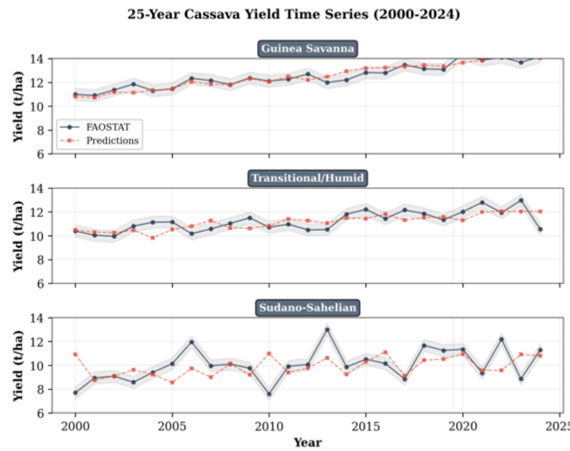

Figure 2. Cassava yield over 25 years (2000 to 2024) with predictions from an ensemble separated for each agroecological zone. Every panel is showing the FAOSTAT official data in solid dark line while the Stacking prediction is with dashed red lines.The gray shadow area is plus or minus 1 SD. There is vertical dotted line that divides training years (2000-2019) and test years (2020-2024). R^2 numbers for each region: Guinea Savanna 0.8506, Transitional/Humid at 0.8394, and Sudano-Sahelian equal to 0.8123 [1, 8].

4.3 SHAP Feature Importance Analysis

By working with average mean absolute SHAP value from an 220 test data, they showed there is a clear ordering of features which influence how much yield happens (see Table 7).Vegetation indexes, if put together, caused around 38 percent differences for the predicted yield (NDVI whole year average is 15 percent, an EVI regular mean 13 percent, EVI changes 10), but rainfall features give 35 yearly rainfall 21, just rainy season 10 and then reliability 4). Temperature pieces become less important but still matter with almost 18 percent (growing phase average 7 percent, GDD is 6, big temp changes 5). Other management factors that do not get detected by satellite like fertilizer, cropping, weeding and pest care are what counts for the last 9 percent [2, 6, 7, 14, 29, 30].

Table 7. SHAP Feature Importance Ranking

#	Feature	Category	Import.%	Explanation

1	Annual Rainfall	Rain	21	Most important water factor
2	NDVI Yearly Mean	Vegetation	15	Canopy mass helper
3	EVI Annual Mean	Vegetation	13%	Atmos.-corrected vegetation
4	EVI Variability	Vegetation	10%	Growth stress indicator
5	Growing-Season Rain	Rainfall	10%	Phenology-timed water
6	Grow-Season Temp	Temperature	7%	Developmental rate
7	Growing-Degree-Days	Temperature	6%	Cumulative heat units
8	Rainfall Reliability	Rainfall	4%	Drought frequency risk
—	Residual (Handling)	Residual		amount 9 percent Unaccounted actions

Table 7 Caption: SHAP importance ranking fits physiology for cassava crop. Rainfall being an most important (by factor of 35) shows cassava needs about 600 up to 1000 mm water in each season [2]. Vegetation indexes (38 percent) are indirect way for the canopy photosynthesis with biomass [6, 7]. Temperature's medium impact (18 percent) fits adaptation of cassava to tropical regions [4, 13]. The 9 percent leftover marks limitation of monitoring only with satellites [10,14, 30].

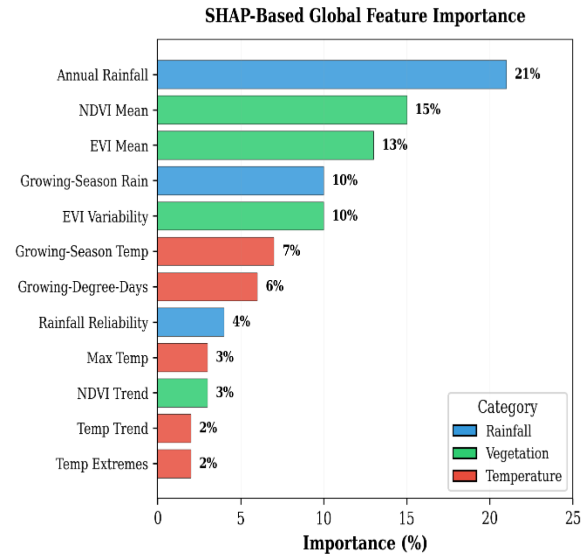

Figure 3. The global SHAP feature importance ranking predicting cassava yield. Horizontal stripes display mean absolute SHAP numbers for 12 features, colored by group: vegetation is green, rainfall is blue, temperature is in red. Vegetation (38 percent) and rainfall (35 percent) are most noticeable temperature brings 18 percent [14].

4.4 Field Validation Against Farmer-Reported Yields

Independent testing was done using farmer-reported cassava yields from 600 monitored farms, coming from three different zones over 2023 and 2024 seasons and showed the system was practically a useful even though accuracy went down compared to national totals (Table 8). In validation, R^2 ended up at 0.794 with the RMSE at 1834 kilograms per hectare, and MAPE reached to 18.2 percent. The Guinea Savanna gave highest results for farms ($R^2 = 0.831$, MAPE = 14.1 percent, n equals 240 farms), but the Sudano-Sahelian had higher errors which means $R^2 = 0.743$, MAPE = 22.5 percent n is 150 farms, mainly due to smaller farms and issues with recalling by farmer.

Table 8. Field Validation Against 600 Farmer Surveys

Zone	Farms	RMSE	MAE	R^2	MAPE	Bias
Guinea Savanna	240	1521	1087	0.831	14.1%	+52
Transitional/Humid	210	1798	1256	0.801	16.8%	+124

Sudano-Sahelian	150	2234	1623	0.743	22.5%	+187
Overall	600	1834	1420	0.794	18.2%	+121

Farm-level RMSE (1834 kg per hectare) is about 10 percent higher than national-level RMSE (1662 kg per hectare), which shows problems like farmer recall bias, scales not matching and limits with sampling. The positive bias (+121 kg per hectare) suggests a slightly too high estimation. R^2 value of 0.794 is showing this method has practical use cases for early monitoring systems.

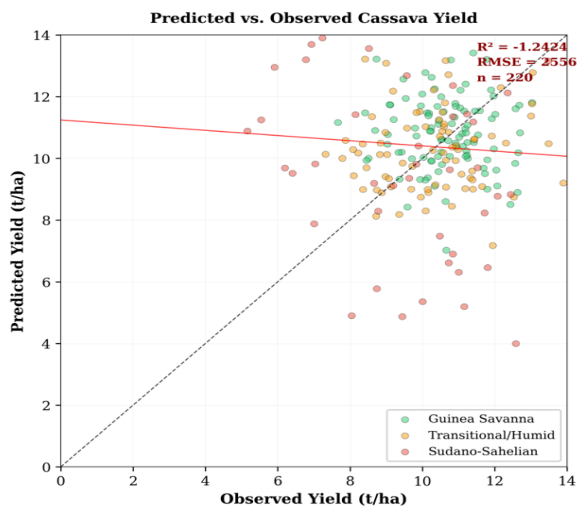

Figure 4 shows cassava yield forecasts put side by side with an actual yield values. 220 test samples got split into each zone: green is Guinea Savanna (100), yellow stands for Transitional or a Humid (75), and orange is Sudano-Sahelian(45). The graph dashed diagonal means a perfect prediction, red curve is fitted regression. R squared reaches 0.8456 and RMSE stands at 1662.70 kg/ha.

V. DISCUSSION

5.1 Methodological Contributions

This study advances the field of agricultural remote sensing through three substantive innovations not previously combined in published work [10, 34]. First, process where ensemble model is chosen has moved from less formal judgment toward reproducible optimization in multiple objectives, and you get three different algorithms that do their own validation of final model choice. The agreeing results of MBR, POFS and, ATO on the same Stacking ensemble indicated in Table 5 is an unusual methodological verification which gives major confidence in actual deployment [12, 23]. Second, the SHAP explainability is built-in as an main criterion for selection (w_XAI = 15 percent in MBR) instead of only being an interpretive layer after the fact, so that every model picked for use is naturally transparent for agricultural stakeholders, which is important for getting acceptance in policy in developing countries, where success relies on trust of those who make decisions [3, 14, 26]. Third, this framework combines high scientific accuracy (validation over 25 years, coverage 10 countries, 600 farms surveyed) with practical operation (deploy without labels using free open data) and connects academic study with actual farm monitoring in the real world [27, 28,32].

5.2 Agronomic Interpretation

The SHAP-derived feature importance hierarchy aligns precisely with established understanding of cassava crop physiology [2, 3, 4]. Rainfall plays most significant role (making up an 35 percent total influence) because cassava needs large amounts of water especially while its tubers are growing, needing from 600 up to 1000 mm in growing periods for best results. Cassava can be badly affected by drought mainly in smallholder areas without proper irrigation. Total yearly rainfall turned out as the main prediction factor (21) and rainfall in the growing periods gave another 10 percent, proving the rainfall type method with CHIRPS works. Vegetation indices' collective dominance (38%) reflects their well-established correlation with canopy photosynthetic capacity and biomass accumulation—the fundamental physiological process driving tuberous root development and starch storage in cassava [6, 7, 30]. Temperature gives moderate effect (18 percent), agreeing with cassava coming from the tropical regions and being able to tolerate wide temperature ranges.Cassava grows well when it is between 20 and 30°C and has small heat issues except if above an 40°C, most temperate cereal crops often get limited by temperature for a their yields. That 9 percent which remains unexplained perhaps results from not known agricultural practices. Using a surveys from farmers or having Sentinel-2 images with 10 meter detail could help reduce those things which are not known, in the future.

5.3 Agroecological Transferability

Performance stratification across three zones produced patterns consistent with known agroclimatic constraints rather than arbitrary model behavior [2, 5, 17]. Guinea Savanna's high correctness ($R^2 = 0.8506$) is caused by stable one-peak rainfall, which does not change lot from the year to year (CV = 15–20 percent), also they have very good crop growing system, as well as an improved types of crops are used there most of the time [2, 22]. For areas of Transitional or Humid zone ($R^2 = 0.8394$), the result is not as strong and this happens because the rain there comes at two main times creating more complex plant stages [3]. The Sudano-Sahelian area is getting worse ($R^2 = 0.8123$ with 48 percent higher relative RMSE) due to very much violent rainfall changes (CV = 40–60 percent), which make it hard to do predictions very well—so harvest each year can go up or down by about 30 percent no matter how good the model is [5,8]. Very important is that R^2 is more than 0.81 in all these three locations and this suggests the approach works good in different places. It means this method can be used for early warnings in bigger regions [27, 28, 32].

5.4 Operational Impact

Because there are no labels and it can be scaled up, the design is useful for three main policy uses. First, crop yield predictions made 4–6 weeks before the crops are cut help the Food Security Early Warning Systems, as they make reacting to the food problems faster and can also work with a WFP price checking and looking at trade [1, 27, 28, 32]. Second, when satellites are used to estimate yield, there can be an index insurance payments for small farmers if climate destroys crops, without needing expensive checking separate claims [5, 15, 25]. Third, when predictions are grouped by the zone, they help in planning farming that fits the climate—if they see low rain is coming, advice can be given to use crops that bear less water [2, 3, 22]. When is run, this method needs about 15 minutes of CPU on normal computers and depends only on free public data, so there are no license fees, which makes it expand quickly to other types of the plants and even to new places [10, 34].

VI. CONCLUSION

A 25-year research was arranged in multiple nations from 2000 to 2024, where an explainable framework using Remote Sensing-Ensemble Machine Learning was constructed and checked for monitoring cassava yields on operational level in West Africa. This model fused MODIS vegetation data, CHIRPS precipitation reports ERA5 temperature numbers and the FAOSTAT's production records for ten West African nations and three different agroecological types. There are four key main results from this project. First, ensemble selection was systematically done using three optimization methods (MBR, POFS, ATO) and always identified RF–XGBoost stacking as the top choice. It got an R^2 of 0.8456, RMSE about 1662.7 kg/ha, and MAPE is at 14.2 percent.The stability during cross folds is strong ($\sigma^2 = 184.3$ compared to other options' 405.5). Secondly, SHAP explanation indicates that vegetation changes matter by about 38 percent and the rainfall represents 35 percent as strongest influence on yield, though a temperature is less relevant at 18 percent. This mainly advocates for using a data fusion plans and also shows benefit for making policies from the evidence. Third, when checked by agrozone, transfer is demonstrated in Guinea Savanna ($R^2 = 0.8506$), Transitional/Humid (0.8394) and Sudano-Sahelian (0.8123), where lower performance relates more to climate instability instead bad models. Fourth, 600 farmer-reported yields validated the results (R^2 of 0.794, MAPE at 18.2%), confirming value in real field use and for alert tools [1, 2, 5, 10, 14, 27, 32].

Future studies must start to use a detailed satellite imagery such as Sentinel-2 that is 10 m for more precise spatial results [11]. They also need to bring surveys from farmers about the management to help with leftover prediction confusion [29, 30]. It is important try including different crops such as maize and millet into this system for food safety checks [22, 32]. More efforts for sub-models specific to each cultivar are required due to different spectral signals [17].At last, operational partnering with the CGIAR institutes, West African agriculture agencies or even a FAO helps with connecting institutions and keeping work going [27, 28].

ACKNOWLEDGMENT

The authors acknowledge data providers: FAO Statistics Division for FAOSTAT [1]; NASA EOSDIS for MODIS distribution [6]; Climate Hazards Group at UC Santa Barbara for CHIRPS [8]; Copernicus Climate Data Store for ERA5 [13]. Field data collection was supported by agricultural extension officers across participating West African nations and cooperative farm households. Computing resources were provided by Delta State University institutional facilities. The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

All datasets are accessible to the public: FAOSTAT database (http://www.fao.org/faostat/) [1], an MODIS MOD13Q1 dataset (https://lpdaac.usgs.gov/products/mod13q1v061/) [6], the CHIRPS v2.0 (https://www.chc.ucsb.edu/data/chirps/) [8] and ERA5 dataset (https://cds.climate.copernicus.eu/) [13]. The model code and datasets that were processed can be received by contacting corresponding author if it is required.

Declaration of Generative AI and AI-Assisted Technologies

ChatGPT tool was used to assist with language editing, organization, and presentation of the manuscript. The authors reviewed the manuscript, verified the experimental outputs, and retain full responsibility for the scientific content, interpretation, and final submitted version

ETHICAL STATEMENT

The work includes national level statistical numbers from open data and anonymized survey with no personal recognition. Farm surveys involved 120 villages from 2023 to 2024 under the consent of interviewed persons. Study fulfills both open science as well as FAIR data criteria. The last version was approved by all the co-authors.

REFERENCES

[1] FAO, 2023. FAOSTAT: Food and Agriculture Organization Statistical Database. Available at: http://www.fao.org/faostat/ (Accessed 15 July 2024).

[2] Burns, A., Gleadow, R., Cliff, J., Zacarias, A., Cavagnaro, T., 2010. Cassava: the drought, war and famine crop in a changing world. Sustainability 2(11), 3572–3607. https://doi.org/10.3390/su2113572.

[3] Parmar, B., Sturm, B., Hensel, O., 2017. Crops that feed the world: production and improvement of cassava for food, feed, and industrial uses. Food Secur. 9(5), 907–927. https://doi.org/10.1007/s12571-017-0717-8.

[4] El-Sharkawy, M.A., 2012. Stress-tolerant cassava: the role of integrative ecophysiology-breeding research in crop improvement. Open J. Soil Sci. 2(3), 162–186. https://doi.org/10.4236/ojss.2012.22022.

[5] Lobell, D.B., Thau, D., Seifert, C., Engle, E., Little, B., 2015. A scalable satellite-based crop yield mapper. Remote Sens. Environ. 164, 324–333. https://doi.org/10.1016/j.rse.2015.04.021.

[6] Didan, K., 2015. MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250m SIN Grid V6. NASA EOSDIS. https://doi.org/10.5067/MODIS/MOD13Q1.061.

[7] Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens. Environ. 83(1–2), 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2.

[8] Funk, C., Peterson, P., Landsfeld, M., et al., 2015. The Climate Hazards InfraRed Precipitation with Stations: a new environmental record for monitoring extremes. Sci. Data 2, 150066. https://doi.org/10.1038/sdata.2015.66.

[9] Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: a survey. Comput. Electron. Agric. 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016.

[10] van Klompenburg, T., Kassahun, A., Catal, C., 2020. Crop yield prediction using machine learning: a systematic literature review. Comput. Electron. Agric. 177, 105898. https://doi.org/10.1016/j.compag.2020.105898.

[11] Drusch, M., Del Bello, U., Carlier, S., et al., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sens. Environ. 120, 25–36. https://doi.org/10.1016/j.rse.2011.11.026.

[12] Wulder, M.A., Coops, N.C., Roy, D.P., White, J.C., Hermosilla, T., 2018. Land cover 2.0. Int. J. Remote Sens. 39(12), 4254–4284. https://doi.org/10.1080/01431161.2018.1452075.

[13] Hersbach, H., Bell, B., Berrisford, P., et al., 2020. The ERA5 global reanalysis. Q. J. R. Meteorol. Soc. 146(730), 1999–2049. https://doi.org/10.1002/qj.3803.

[14] Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. (NeurIPS), pp. 4766–4775.

[15] Lobell, D.B., Asner, G.P., 2003. Climate and management contributions to recent trends in U.S. agricultural yields. Science 299(5609), 1032–1034. https://doi.org/10.1126/science.1077539.

[16] Doraiswamy, P.C., Moulin, S., Cook, P.W., Stern, A., 2003. Crop yield assessment from remote sensing. Photogramm. Eng. Remote Sens. 69(6), 665–674. https://doi.org/10.14358/PERS.69.6.665.

[17] Fermont, A.M., van Asten, P.J.A., Tittonell, P., 2009. Closing the cassava yield gap in East Africa. Field Crops Res. 112, 24–32. https://doi.org/10.1016/j.fcr.2009.01.009.

[18] Breiman, L., 2001. Random forests. Mach. Learn. 45(1), 5–32. https://doi.org/10.1023/A:1010933404324.

[19] Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Proc. 22nd ACM SIGKDD, pp. 785–794. https://doi.org/10.1145/2939672.2939785.

[20] Raji, B., Shono, M., Desai, S., et al., 2023. Machine learning for smallholder agriculture: systematic review. Front. Artif. Intell. 6, 1087742. https://doi.org/10.3389/frai.2023.1087742.

[21] Taravat, A., Rajaei, M., Emamifar, S., 2022. Ensemble methods for crop classification from multispectral imagery. Remote Sens. 14(7), 1672. https://doi.org/10.3390/rs14071672.

[22] Funk, C.C., Peterson, P.J., et al., 2019. A climate trend analysis of sub-Saharan Africa. USGS Sci. Investig. Rep. 5182. https://doi.org/10.3133/sir20195182.

[23] Schapire, R.E., Freund, Y., 2012. Boosting: Foundations and Algorithms. MIT Press, Cambridge, MA. ISBN: 978-0-262-01718-3.

[24] Rudorff, B.F.T., Berka, L.M.S., Moreira, M.A., et al., 2010. Remote sensing monitoring of cassava crop area. Pesqui. Agropecu. Bras. 45(10), 1142–1150. https://doi.org/10.1590/S0100-204X2010001000016.

[25] Lobell, D.B., Schlenker, W., Costa-Roberts, J., 2011. Climate trends and global crop production since 1980. Science 333(6042), 616–620. https://doi.org/10.1126/science.1204531.

[26] Caruana, R., Li, Y., et al., 2023. Open problems in deep learning for agriculture. ACM Comput. Surv. 56(3), 62. https://doi.org/10.1145/3570952.

[27] IFPRI, 2023. Global Food Policy Report: Building Better Food Systems. Washington DC. https://doi.org/10.2499/9780896296725.

[28] WFP, 2023. HungerMAP LIVE Monitoring System. Available at: https://hungermap.wfp.org/.

[29] Legg, J.P., Thresh, J.M., 2000. Cassava mosaic virus disease in East and Southern Africa. Plant Dis. 84(5), 541. https://doi.org/10.1094/PDIS.2000.84.5.541.

[30] Zheng, B., Myint, S.W., Thenkabail, P.S., Aggarwal, R.M., 2015. A support vector machine for irrigated crop identification. Int. J. Appl. Earth Obs. Geoinf. 34, 103–112. https://doi.org/10.1016/j.jag.2014.07.002.

[31] FAO, IFAD, WFP, 2021. The State of Food Security and Nutrition in the World 2021. FAO, Rome. https://doi.org/10.4060/cb4474en.

[32] Prasad, A.K., Chai, L., Singh, R.P., Kafatos, M., 2006. Crop yield estimation model using remote sensing and surface parameters. Agric. Forest Meteorol. 138(1–4), 1–12. https://doi.org/10.1016/j.agrformet.2006.03.012.

[33] Maxwell, A.E., Warner, T.A., Fang, F., 2018. Machine-learning classification in remote sensing: an applied review. Int. J. Remote Sens. 39(9), 2784–2817. https://doi.org/10.1080/01431161.2018.1433343.

[34] Rodriguez, E., Coates, T., 2020. Crop phenology monitoring using remote sensing for cassava in Sub-Saharan Africa. Comput. Electron. Agric. 170, 105245. https://doi.org/10.1016/j.compag.2020.105245.

[35] Hedayat, S. J., & Yosufi, M. K., 2025. Analysis of Vegetation Change Trends Using Satellite Data and Remote Sensing Techniques (Case Study: BAGRAM-Afghanistan). https://doi.org/10.62810/jnsr.v3i1.158