agriculture

MDPI

Article

Field-Scale Rice Yield Prediction in Northern Coastal Region of Peru Using Sentinel-2 Vegetation Indices and Machine Learning Models

Isabel Jarro-Espinal [1], José Huanuqueño-Murillo [2], Javier Quille-Mamani [3], David Quispe-Tito [2], Lia Ramos-Fernández [2,*], Edwin Pino-Vargas [4] and Alfonso Torres-Rua [5]

[1] Doctoral Program in Water Resources, Graduate School, National Agrarian University La Molina, Lima 15024, Peru; 20211518@lamolina.edu.pe
[2] Departament of Water Resources, National Agrarian University La Molina, Lima 15024, Peru; 20241699@lamolina.edu.pe (J.H.-M.); 20211471@lamolina.edu.pe (D.Q.-T.)
[3] Geo-Environmental Cartography and Remote Sensing Group (CGAT), Universitat Politècnica de València, Camí de Vera s/n, 46022 Valencia, Spain; jaquille@upv.es
[4] Departament of Civil Engineering, Jorge Basadre Grohmann National University, Tacna 23000, Peru; epinov@unjbg.edu.pe
[5] Civil and Environmental Engineering Department, Utah State University, Old Main Hill, Logan, UT 84322, USA; alfonso.torres@usu.edu
* Correspondence: liarf@lamolina.edu.pe

Abstract

Accurate rice yield prediction is essential for optimizing water management and supporting decision-making in agricultural systems, particularly in arid environments where irrigation efficiency is critical. This study assessed five machine learning algorithms—Multiple Linear Regression (MLR), Support Vector Regression (SVR, linear and RBF), Partial Least Squares Regression (PLSR), Random Forest (RF), and Extreme Gradient Boosting (XGBoost)—for plot-scale rice yield estimation using Sentinel-2 vegetation indices (VIs) during the 2022 and 2023 seasons in the Chancay–Lambayeque Valley, Peru. VIs sensitive to canopy vigor, water status, and structure were derived in Google Earth Engine and optimized via Sequential Forward Selection to identify the most relevant predictors per phenological stage. Models were trained and validated against field yields using leave-one-out cross-validation (LOOCV). Intermediate stages (Flowering, Milk, Dough) yielded the strongest relationships, with water-sensitive indices (NDMI, MSI) consistently ranked as key predictors. MLR and PLSR achieved the highest generalization (R^2_CV up to 0.68; RMSE_CV $\approx$ 1.3 t ha^{-1}), while RF and XGBoost showed high training accuracy but lower validation performance, indicating overfitting. Model accuracy decreased in 2023 due to climatic variability and limited satellite observations. Findings confirm that Sentinel-2–based VI modeling offers a cost-effective, scalable alternative to UAV data for operational rice yield monitoring, supporting water resource management and decision-making in data-scarce agricultural regions.

Keywords: Sentinel-2; vegetation indices (VIs); rice yield prediction; machine learning models; multiple linear regression (MLR); support vector regression (SVR); partial least squares regression (PLSR); random forest (RF); extreme gradient boosting (XBoost); cross-validation (CV)

check for updates

Academic Editor: Jiehao Li

Received: 3 September 2025
Revised: 26 September 2025
Accepted: 29 September 2025
Published: 30 September 2025

Citation: Jarro-Espinal, I.; Huanuqueño-Murillo, J.; Quille-Mamani, J.; Quispe-Tito, D.; Ramos-Fernández, L.; Pino-Vargas, E.; Torres-Rua, A. Field-Scale Rice Yield Prediction in Northern Coastal Region of Peru Using Sentinel-2 Vegetation Indices and Machine Learning Models. *Agriculture* **2025**, *15*, 2054. https://doi.org/10.3390/agriculture15192054

Correction Statement: This article has been republished with a minor change. The change does not affect the scientific content of the article and further details are available within the backmatter of the website version of this article.

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Rice (*Oryza sativa* L.) is a staple food for more than half of the global population and plays a crucial role in global food security [1]. In the Chancay–Lambayeque basin (Peru),

this crop occupies a significant proportion of the irrigated agricultural area, contributing substantially to the local economy and rural livelihoods. However, rice production also exerts considerable pressure on water resources, as its cultivation may require between 3000 and 5000 L of water per kilogram of grain produced [2]. In this context, accurate yield prediction methods are essential not only for improving production efficiency but also for optimizing regional water-use planning [3,4].

Satellite remote sensing has revolutionized precision agriculture by enabling continuous, spatially detailed monitoring of crop growth and development. Among the available platforms, Sentinel-2 stands out for its combination of 13 spectral bands spanning the visible, red-edge, near-infrared, and shortwave infrared regions, with spatial resolutions of 10–20 m and a revisit time of five days at the equator. These characteristics allow the generation of dense time series that capture phenological variations throughout the crop cycle. Sentinel-2–derived vegetation indices (VIs), such as NDVI, EVI, NDMI, and MSI, have proven to be effective tools for assessing canopy vigor, water status, and biomass accumulation [4–6]. Nonetheless, cloud cover during critical phenological stages and the relatively small plot sizes in smallholder farming systems can limit retrieval accuracy [1].

Machine learning (ML) has emerged as an effective tool for exploiting the high dimensionality of remote sensing data in agricultural yield prediction. These methods can model complex, non-linear relationships between spectral variables and yield indicators, often outperforming traditional statistical approaches. In this study, Multiple Linear Regression (MLR) is used as a baseline model due to its simplicity and competitiveness when predictor collinearity is controlled. Support Vector Regression (SVR), in its linear and radial basis function (RBF) forms, projects input variables into higher-dimensional spaces, enabling both linear and non-linear fitting. Partial Least Squares Regression (PLSR) is particularly suitable when predictors are highly correlated, as it projects them into latent components that maximize covariance with the response variable. Random Forest (RF), based on an ensemble of decision trees, handles complex interactions and provides measures of variable importance, whereas Extreme Gradient Boosting (XGBoost) builds models sequentially, optimizing residual errors and often achieving high accuracy in tabular datasets. Each of these algorithms presents advantages and limitations in terms of interpretability, computational cost, and overfitting risk, which justifies their comparative evaluation [7–9].

Recent studies have highlighted the potential of combining satellite-derived vegetation indices with machine learning algorithms to predict rice yield without relying on UAV platforms. For instance, Islam et al. [1] integrated Sentinel-2 NDVI with meteorological variables (precipitation, soil moisture, and evapotranspiration) using ensemble regression models, achieving high prediction accuracy in Nepal. Similarly, Chen et al. [10] demonstrated that combining Sentinel-2 optical data with Sentinel-1 SAR through multiple regression improves yield and forage quality estimation compared to single-sensor approaches. In the Peruvian context, Quille-Mamani et al. [4] used Sentinel-2 NDVI time series and phenological metrics to predict rice yield with consistent results across seasons. Other works have shown that PLSR [9], SVR and RF [8], and XGBoost [7] can effectively capture complex relationships between vegetation indices and yield, supporting their inclusion in multi-model evaluations.

In this context, the present study exclusively employs Sentinel-2 imagery from the 2022 and 2023 growing seasons to compute a diverse set of vegetation indices, analyzed both individually and in combination across phenological stages. The MLR, SVR, PLSR, RF, and XGBoost models were implemented, optimized through Sequential Forward Selection (SFS), and evaluated using Leave-One-Out Cross-Validation (LOOCV). The objective is to develop robust and transferable models for regional-scale rice yield prediction, providing tools for agricultural planning without reliance on UAV-based observations.

2. Materials and Methods

2.1. Study Sites

The research was conducted in the Lambayeque region, in northwestern Peru, focusing on intensive cultivation areas dedicated to commercial rice production. The study area—located in the province of Ferreñafe—encompasses five agricultural zones: García, Santa Julia, Totora, Zapote, and Caballito, within the Chancay–Lambayeque Valley. A total of 37 fields were monitored during the 2022 growing season and 35 fields in 2023 (see Table 1).

Table 1. Characteristics of the study area.

Zones	Longitude	Latitude	Altitude (m.a.s.l.)	Area (ha)	Sub-plots	Variety
Caballito	06°35′38.82″ S	79°47′5.32″ W	47	14.19	15	Tinajone and Capoteña
García	06°35′2.51″ S	79°47′3.50″ W	47	5.23	3	Tinajones
Santa Julia	06°36′25.99″ S	79°47′31.85″ W	42	8.55	7	Mallares
Totora	06°35′35.16″ S	79°47′32.74″ W	44	5.38	6	Puntilla
Zapote	06°35′44.20″ S	79°47′8.04″ W	46	6.01	6	Pakamuros

According to the Köppen–Geiger climate classification, the area has a hot desert climate (BWh), characterized by mean annual temperatures of approximately 21–22 °C and extremely low annual precipitation, ranging from about 180 to 210 mm [11]. Rice production in this valley relies heavily on irrigation supplied by the Chancay–Lambayeque River system due to the pronounced aridity of the environment. The geographic location of the study area is shown in Figure 1, including its position within Peru and the spatial distribution of the monitored fields across the selected agricultural zones.

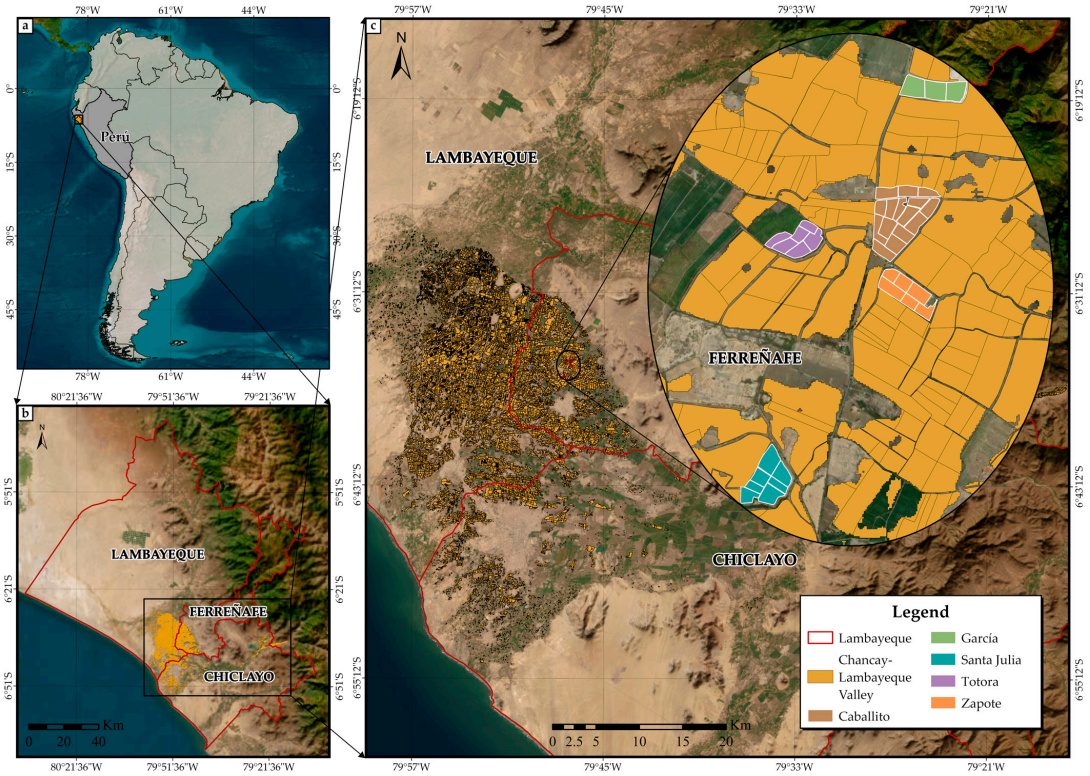


Figure 1. Geographic location of the study area in the Lambayeque region, Peru: (**a**) Location of Lambayeque within Peru, highlighting its position in South America; (**b**) The Chancay–Lambayeque Valley within Lambayeque, showing rice field distribution across the provinces of Lambayeque, Ferreñafe, and Chiclayo; (**c**) Detailed view of monitored rice fields in the districts of Ferreñafe and Chongoyape, including the five selected agricultural zones: García, Santa Julia, Totora, Zapote, and Caballito.

2.2. Meteorological Characteritics

The meteorological conditions in the study area during the rice growing period (January–June) for the 2022 and 2023 seasons are presented in Figure 2. The variables include relative humidity (%), maximum and minimum air temperature (°C), precipitation (mm), and wind speed (m s^{-1}), recorded by the INIA–Vista Florida automatic weather station. Panel (a) shows monthly precipitation and relative humidity, while panel (b) depicts maximum and minimum temperatures together with wind speed trends for both years.

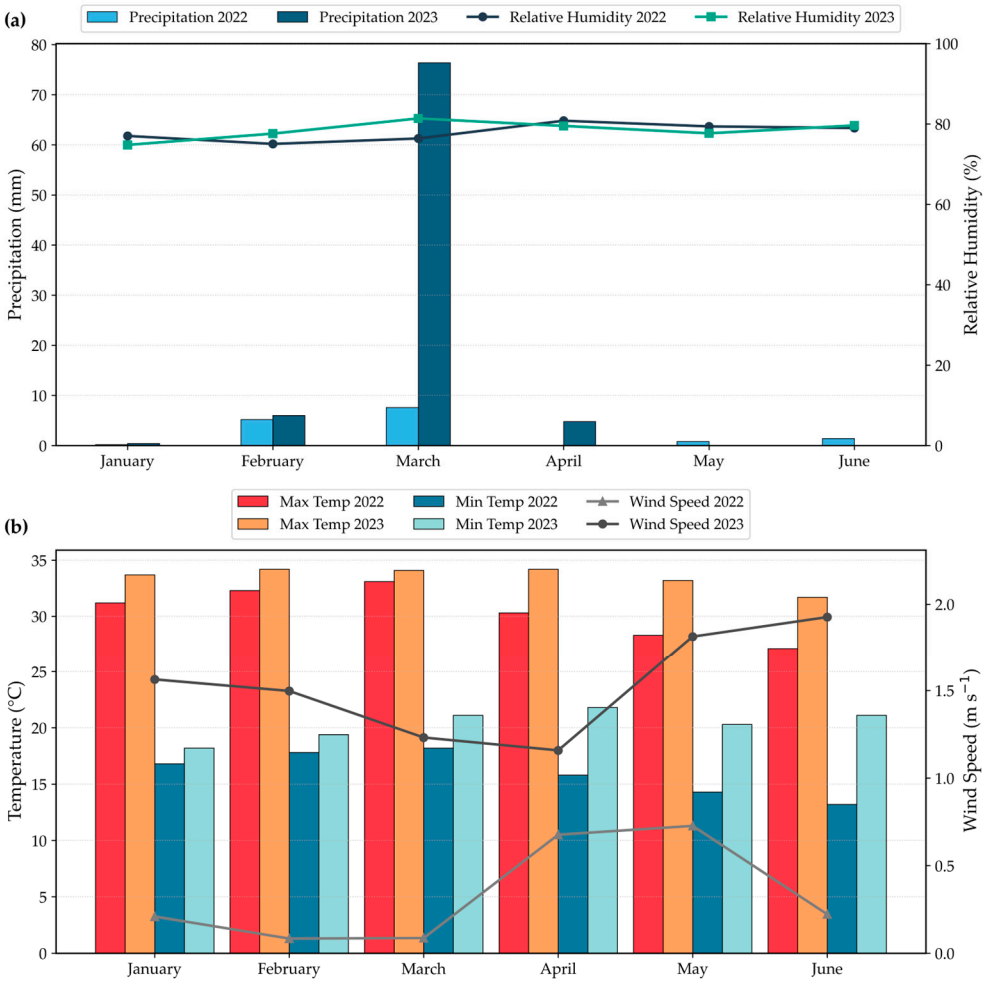


Figure 2. Monthly meteorological variables recorded during the rice growing period (January–June) for the 2022 and 2023 seasons at the INIA–Vista Florida automatic weather station. (**a**) Precipitation (mm) and relative humidity (%). (**b**) Maximum and minimum air temperature (°C) and wind speed (m s^{-1}). These datasets were used to characterize interannual climatic variability in the study area and to support the interpretation of crop development patterns.

2.3. Crop Management and Agronomic Practices

Rice was cultivated during the 2022 and 2023 growing seasons in five commercial zones: Caballito, Totora, García, Santa Julia, and Zapote (Table 1). In 2022, 37 subplots were monitored, whereas in 2023, the number decreased to 35 due to salinity problems. Subplot sizes ranged between 5 and 12 ha. The main cultivars were Tinajones, Capoteña, Puntilla, Mallares, and Pakamuros, which are widely grown in the Chancay–Lambayeque Valley due to their tolerance to water stress and high yield potential. Tinajones (semi-early cycle) typically yields 14–15 t ha^{-1}, Capoteña (intermediate cycle) 12–13 t ha^{-1}, while Puntilla, Mallares, and Pakamuros can reach up to 13 t ha^{-1} under optimal agronomic conditions.

Sowing was performed in seedbeds, and transplanting was conducted manually at 35 days after sowing (DAS), using a planting frame of 25 × 25 cm with two seedlings per hill. Crop phenology was monitored with growing degree days (GDD), which allowed for precise identification of key growth stages and supported the timing of management practices. All commercial plots followed a continuous flooding (CF) irrigation system, ensuring permanent water availability throughout the crop cycle, which is the standard practice in this irrigated desert valley.

Fertilization rates varied slightly across zones: Zapote received 265 N–72 P–50 K, whereas the other areas received 263 N–92 P–75 K. Fertilizer application was split into three stages: before transplanting (5% N, 85% P, 100% K), at 18 DAS (42% N, 15% P), and the remaining nitrogen was applied in equal splits during tillering and panicle initiation (cotton point). Harvesting was carried out mechanically between 140 and 166 DAS in 2022 and 139 and 157 DAS in 2023, depending on the subplot.

2.4. Data Acquisition and Processing

Multispectral Sentinel-2 imagery corresponding to the 2022–2023 period was acquired, including bands in the visible, red-edge, near-infrared, and shortwave infrared regions (B2, B3, B4, B5, B6, B8, B8A, B11). These images, with spatial resolutions ranging from 10 to 20 m, were obtained from the Sentinel-2 mission under the Copernicus program, widely used in agricultural applications and vegetation monitoring due to its global coverage and high radiometric quality [5].

Image processing was carried out in Google Earth Engine (GEE) to compute a suite of VIs sensitive to key attributes of rice crops, such as vegetative vigor, canopy water content, structural characteristics, and biomass. Similar studies have successfully employed indices such as NDVI, EVI, and other Sentinel-2–derived VIs for crop yield prediction [5,12].

To enhance predictive performance and avoid redundancy, a Sequential Forward Selection (SFS) strategy was applied to iteratively select the most relevant indices. This approach optimizes model performance and reduces overfitting, as demonstrated in studies that integrate feature selection prior to machine learning modeling [13].

With this refined set of VIs, machine learning models were developed in Python 3.13—namely MLR, SVR (with linear and RBF kernels), PLSR, XGBoost, and RF—with hyperparameters tuned via grid search to maximize predictive accuracy. This combination of statistical and ensemble learning approaches has been effectively implemented in rice yield prediction using Sentinel-2 data and platforms such as GEE [5,14].

Finally, models were trained and validated using field-measured yield data, applying a Leave-One-Out Cross-Validation (LOOCV) scheme to assess generalization capability. The overall methodological workflow is summarized in Figure 3, highlighting the integration of Sentinel-2 data, field measurements, feature selection, model optimization, and validation steps.

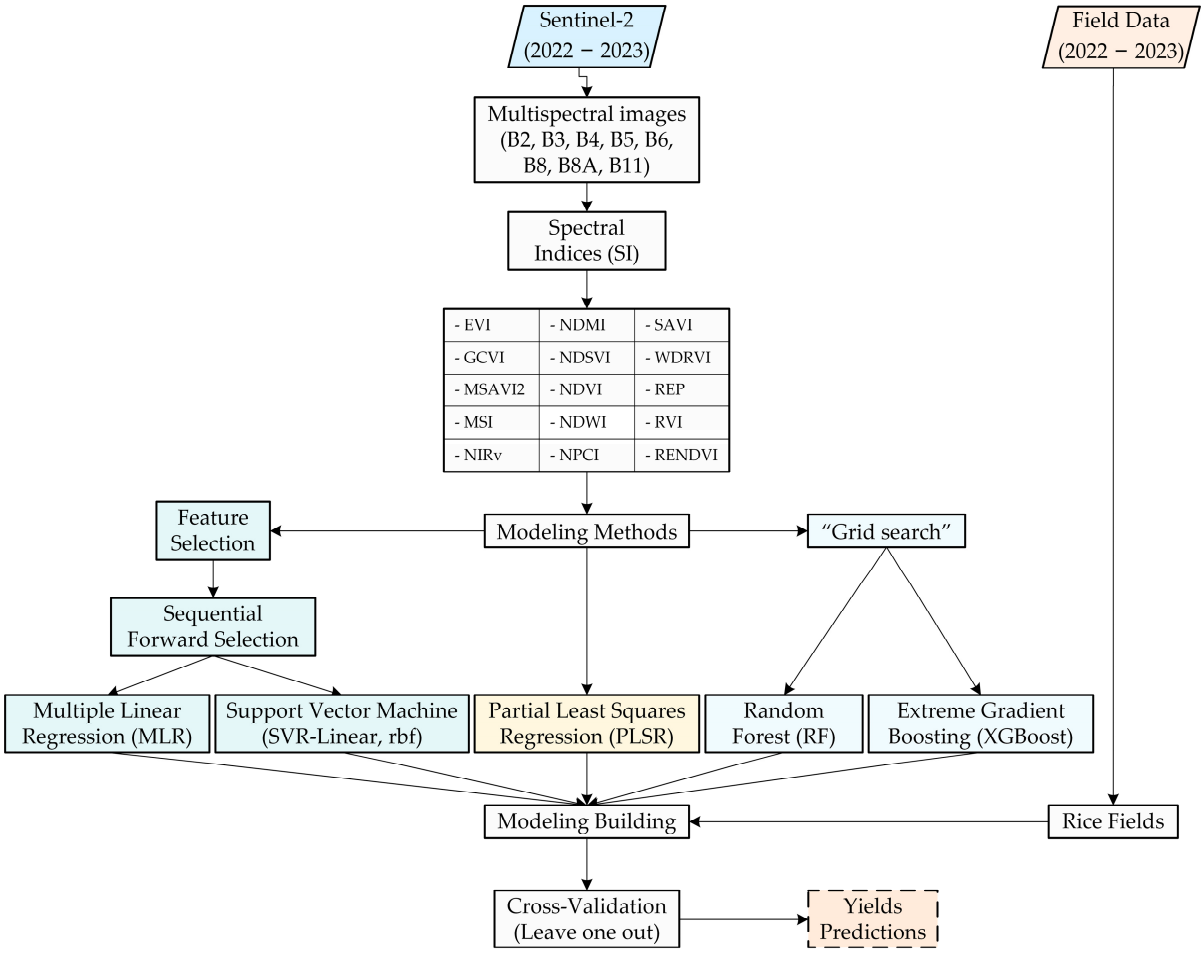


Figure 3. Workflow for rice yield prediction (2022−2023) using Sentinel-2 spectral indices and machine learning models with feature selection and cross-validation.

2.4.1. Sentinel-2 Data Acquisition and Spatial Resolution Harmonization

The spectral characteristics and spatial resolutions of the Sentinel-2 bands used in this study are summarized in Table 2. These include Blue (B2), Green (B3), and Red (B4) bands; two Red-edge bands (B5 and B6); broad Near-Infrared (B8); narrow Near-Infrared (B8A); and two Shortwave Infrared (SWIR) bands (B11 and B12). Bands B2, B3, B4, and B8 have a native spatial resolution of 10 m, whereas B5, B6, B8A, B11, and B12 are acquired at 20 m resolution.

Table 2. Spectral characteristics and spatial resolution of the Sentinel-2 bands used in this study.

Band Name	Sentinel-2 (Band)	Spectral Range (nm)	Resolution (m)
Blue	B2	480–523	10
Green	B3	543–578	10
Red	B4	650–680	10
Red-edge 1	B5	698–713	20
Red-edge 2	B6	733–748	20
Near-infrared (broad)	B8	785–900	10
Near-infrared (narrow)	B8A	855–875	20
SWIR1	B11	1565–1655	20
SWIR2	B12	2100–2280	20

Sentinel-2 imagery was first filtered to retain only scenes with <30% overall cloud cover, applying the Scene Classification Layer (SCL) mask to remove cloud, cirrus, and shadow pixels. This filtering was applied consistently across all phenological stages and both growing seasons (2022 and 2023). After filtering, the 2022 season yielded approximately one-third of the total acquisitions (≈33% usable), with Dough and Maturity having the best coverage, while Flowering and Heading were more affected by cloud contamination. In contrast, the 2023 season was particularly limited (≈12% usable acquisitions), especially during Flowering and Heading, although the Dough stage reached ~33% availability. These interannual differences in data density help explain part of the variability in model performance.

Subsequently, to ensure spatial consistency across all bands and enable reliable interband comparisons, the 20 m Sentinel-2 bands were resampled to 10 m using bicubic interpolation and reprojected to match the spatial reference system of the 10 m bands within the Google Earth Engine (GEE) platform. This harmonization step was essential for pixel-wise spectral index computation and multivariate modeling. Resampling is a standard practice in remote sensing to homogenize spatial resolutions when integrating multiple spectral bands [15]. Bicubic interpolation, in particular, is widely employed in GEE and other platforms because it generates smooth outputs, preserves spatial continuity, and minimizes artifacts compared to simpler techniques such as nearest-neighbor or bilinear interpolation [15–17]. In vegetation monitoring, bicubic interpolation has demonstrated acceptable error margins (e.g., ∼0.19 LAI RMSE) and performance comparable to ESA's operational methods such as SNAP [18]. It has also been successfully applied as a low-frequency component generator in spectral–spatial sharpening frameworks, proving effective for enhancing lower-resolution bands such as SWIR [19]. Moreover, in comparative studies of super-resolution techniques, bicubic interpolation consistently serves as a robust baseline against advanced deep learning models, maintaining adequate accuracy while ensuring computational simplicity [20]. Together, these findings support the appropriateness of bicubic interpolation in harmonizing multispectral Sentinel-2 data for accurate time series reconstruction and vegetation index analysis.

2.4.2. Measurement of Rice Grain Yield

To evaluate rice yield, the process began by determining the moisture content of dehusked grains using a WILE-55 moisture meter (Farmcomp Oy, Espoo, Finland). This device measures moisture within the range of 8–30% with an accuracy of ±0.5%, operating effectively at temperatures between 0 °C and 40 °C. Grain samples were measured when their moisture content reached approximately 16%, and yields were standardized to ~14% moisture during the weighing process to ensure consistency. Harvesting in the commercial fields was carried out mechanically, occurring between 140 and 166 days post-sowing (DPS) in 2022 and between 139 and 157 DPS in 2023, depending on the specific subplot assessed (refer to Figure 4).

The study area was divided into five subareas: Zapote (FPF-Z), Caballito (FSV-CB), García (FSV-G), Santa Julia (FSV-SJ), and Totora (FSV-T). Each plot code (e.g., FSV-CB101) refers to a specific rice plot within these subareas, where the prefix indicates the subarea (e.g., FSV-CB = Caballito) and the numeric suffix identifies the plot number. Figure 4 presents the yield (t ha^{-1}) for each rice plot across the two growing seasons. Blue bars represent the 2022 harvest, while orange bars represent the 2023 harvest. Overall, variability in yield is evident both between years and among plots, with certain plots in Zapote (FPF-Z) and Caballito (FSV-CB) achieving the highest yields in 2022, while some plots in 2023 showed a noticeable reduction, potentially due to differences in environmental conditions, management practices, or other agronomic factors.

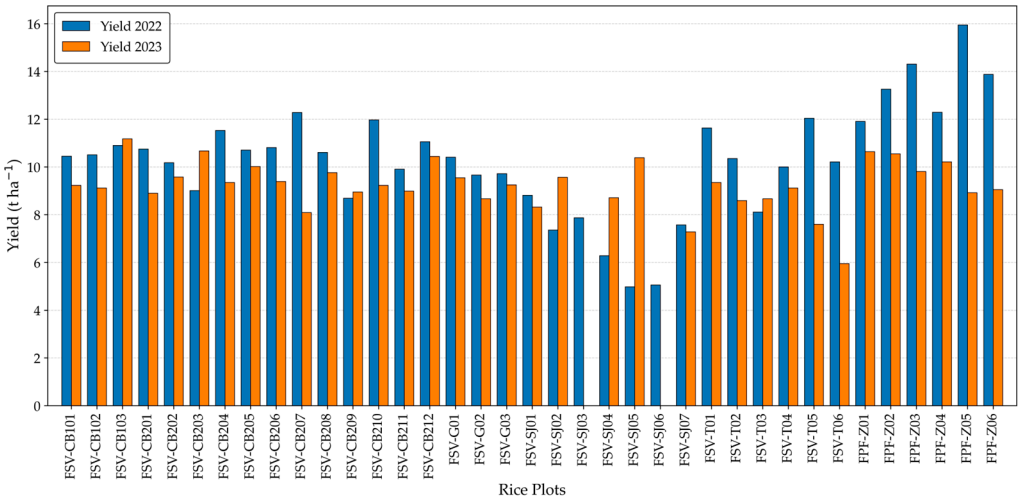


Figure 4. Rice yield (t ha^{-1}) in commercial plots across five subareas of the Chancay–Lambayeque Valley: Zapote (FPF-Z), Caballito (FSV-CB), García (FSV-G), Santa Julia (FSV-SJ), and Totora (FSV-T). Each plot code indicates the subarea prefix and the individual plot number (e.g., FSV-CB101 = Caballito, plot 101).

2.4.3. Sentinel-2 Spectral Indices and Their Computation

The 15 selected spectral indices (Table 3) include both traditional indices such as NDVI, SAVI, and RVI, as well as more recent indices derived from Red Edge (REP, RENDVI) and SWIR bands (NDMI, MSI, NDSVI). Several of these indices were implemented based on the standardized Awesome Spectral Indices (ASI) database provided by the spyndex library [21], which ensures reproducibility and consistency of the formulas employed. The use of ASI allowed for the transparent integration of indices related to canopy vigor, chlorophyll content, and leaf water status, all of which are widely validated in agricultural remote sensing applications.

Table 3. Spectral indices derived from Sentinel-2 imagery, their calculation formulas, and bibliographic references.

Spectral Indices	Calculation Formula	Sources
Normalized Difference Vegetation Index (NDVI)	$NDVI = \frac{NIR-Red}{NIR+Red}$	[4,5,13,22–24]
Enhanced Vegetation Index (EVI)	$EVI = 2.5 \times \frac{NIR-Red}{NIR+6 \times Red-7.5 \times Blue+1}$	[13,22,25]
Soil Adjusted Vegetation Index (SAVI)	$SAVI = 1.5 \times \frac{(NIR-Red)}{NIR+Red+0.5}$	[23,24,26–30]
Modified Soil Adjusted Vegetation Index 2 (MSAVI2)	$MSAVI2 = \frac{2NIR+1-\sqrt{(2NIR+1)^2-8(NIR-Red)}}{2}$	[23,30]
Normalized Difference Moisture Index (NDMI)	$NDMI = \frac{NIR-SWIR1}{NIR+SWIR1}$	[25,31]
Normalized Difference Water Index (NDWI)	$NDWI = \frac{Green-NIR}{Green+NIR}$	[24,30,32–34]
Ratio Vegetation Index (RVI)	$RVI = \frac{NIR}{Red}$	[13,23,28,29]
Moisture Stress Index (MSI)	$MSI = \frac{SWIR}{NIR}$	[25,35]
Red Edge Position (REP)	$REP = \frac{RedEdge1+RedEdge2}{2}$	[9,36]
Red Edge NDVI (RENDVI)	$RENDVI = \frac{NIR-RedEdge}{NIR+RedEdge}$	[36,37]
Green Chlorophyll Vegetation Index (GCVI)	$GCVI = \frac{NIR}{Green} - 1$	[38–40]
Near-Infrared Reflectance of Vegetation (NIRv)	$NIRv = NDVI \times NIR$	[22,41,42]
Normalized Pigment Chlorophyll Ratio Index (NPCI)	$NPCI = \frac{Red-Blue}{Red+Blue}$	[36,43]
Wide Dynamic Range Vegetation Index (WDRVI)	$WDRVI = \frac{\alpha NIR-Red}{\alpha NIR+Red}$	[44–46]
Normalized Difference Senescence Vegetation Index (NDSVI)	$NDSVI = \frac{SWIR-Green}{SWIR+Green}$	[47,48]

2.5. Modeling Methods

2.5.1. Multiple Linear Regression (MLR)

This method is a well-established statistical approach that models the relationship between a dependent variable and a set of independent variables through their linear combination [49]. The method estimates regression coefficients by minimizing the sum

of squared errors, aiming to obtain the best possible linear fit between predictors and the response variable. A key advantage of MLR lies in its simplicity and interpretability, as it allows the individual contribution of each predictor to be quantified [50]. However, its performance may be compromised by multicollinearity and overfitting when the number of predictors is large or when they are highly correlated. In this study, MLR was implemented as a baseline model to benchmark the performance of more complex algorithms, with the mean squared error (MSE) used as the primary evaluation metric.

2.5.2. Support Vector Regression with Sequential Forward Selection (SFS-SVR)

The Sequential Forward Selection (SFS) algorithm begins by using a single feature to build a model with the specified algorithm and iteratively adds the variable that provides the greatest improvement in performance until the optimal combination of predictors is identified [51]. One of its main advantages is the ability to automatically select a subset of truly relevant features, thereby increasing computational efficiency and reducing generalization error by discarding irrelevant or noisy variables, which in turn promotes more accurate and stable models.

When combined with Support Vector Regression (SVR), the process starts with an empty set of predictors and progressively incorporates the variable that yields the highest improvement in model performance, repeating the procedure until the optimal combination is found according to a predefined evaluation function [52]. In this study, the SFS-SVR was implemented using both linear and radial basis function (RBF) kernels, with parameters set to $C = 1$, tolerance = 0.001, and $\varepsilon = 0.1$. The mean squared error (MSE) was employed as the minimization criterion, and resampling techniques were applied to stabilize feature ranking. All variable selection and regression modeling processes were carried out in Python using the Mlxtend and Scikit-learn libraries [4,51].

2.5.3. Partial Least Squares Regression (PLSR)

Partial Least Squares Regression (PLSR) is a latent projection method that simultaneously transforms both the predictor variables and the target variable into a new space of latent components [29]. The process begins by identifying directions in the predictor space that exhibit maximum covariance with the response variable, thereby reducing dimensionality and multicollinearity [53]. A key advantage of PLSR is its ability to handle more variables than observations and to operate effectively with highly correlated datasets, making it particularly valuable in agricultural remote sensing where multispectral and hyperspectral data often generate a large number of redundant predictors. In this study, PLSR was used to estimate biophysical variables from spectral indices, with performance evaluated using the mean squared error (MSE) and cross-validation.

2.5.4. Random Forest (RF)

The Random Forest (RF) algorithm is an ensemble learning method that combines multiple decision trees built from bootstrap samples of the data and random subsets of predictor variables, thereby reducing correlation between trees and improving generalization capability [54]. Each tree is constructed using the bootstrap method, and at each node, a random subset of predictors is selected to determine the optimal split [50]. A key advantage of RF is its robustness to noisy data and resistance to multicollinearity, while maintaining high predictive accuracy even with a large number of variables. Additionally, RF provides internal variable importance metrics that help interpret the relative contribution of each predictor to the model [55]. In recent agricultural applications, RF has achieved strong performance in crop yield estimation, often outperforming linear models and other ensemble methods, particularly when integrating multispectral, meteorological, and management data [56,57].

In this study, key hyperparameters such as the number of trees (1–150), maximum tree depth (3–20), and maximum number of predictors per split (2–16) were optimized through a grid search approach. Model performance was assessed using the mean squared error (MSE) and cross-validation. The model was implemented in Python using the Scikit-learn library.

2.5.5. Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) is a boosting algorithm that sequentially trains decision trees, where each new tree corrects the errors of the previous ensemble [58]. The process begins by fitting a base model, and in subsequent iterations, the objective function is optimized via gradient descent, incorporating regularization terms to prevent overfitting [7]. A notable advantage of XGBoost is its high computational efficiency and ability to handle large datasets and high-dimensional feature spaces, while maintaining strong predictive performance.

In this study, hyperparameters such as maximum tree depth, number of trees, and learning rate were optimized using a grid search approach. Model evaluation was performed using the mean squared error (MSE) and cross-validation. The XGBoost model was developed in Python using the XGBoost and Scikit-learn libraries.

2.6. Predictive Accuracy Assesment

A total of 37 field plots were monitored in 2022 and 35 in 2023, resulting in 72 observations across the two growing seasons. Given this relatively limited sample size, model calibration and validation were conducted through exhaustive resampling strategies. For regression models (MLR, SVR, PLSR, XGBoost), Leave-One-Out Cross-Validation (LOOCV) was applied, ensuring that each field plot was iteratively excluded for independent testing. In the case of ensemble models (RF), the out-of-bag (OOB) error was additionally evaluated as an internal validation metric. These approaches reduce potential bias caused by small sample sizes and provide a more robust assessment of model generalization capability.

The predictive accuracy of rice yield models at the plot level was evaluated using three key indicators: the coefficient of determination (R^2), the root mean square error (RMSE), and the relative root mean square error (rRMSE). The R^2 quantifies the proportion of variance in the observed data explained by the model; the RMSE measures the average magnitude of prediction errors in absolute terms; and the rRMSE expresses this error relative to the mean observed value, as a percentage. The combined use of R^2, RMSE, rRMSE, and LOOCV is well established in studies that apply remote sensing and machine learning for crop yield prediction, including approaches based on spectral indices, multi-source data, and ML algorithms [4,59,60]. Together, these procedures provide a comprehensive evaluation of the predictive accuracy of the rice yield models.

The metrics were calculated as follows:

$$R^2 = 1 - \frac{\sum_{j=1}^{n}\left(Y_j - \hat{Y}_j\right)^2}{\sum_{j=1}^{n}\left(Y_j - \bar{Y}\right)^2} \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{j=1}^{n}\left(Y_j - \hat{Y}_j\right)^2}{n}} \quad (2)$$

$$rRMSE = \frac{RMSE}{\bar{Y}} \times 100\% \quad (3)$$

where n is the total number of plots; Y_j and $\hat{Y}_j$ denote the *j*th measured and predicted values of rice yield, respectively; $\overline{Y}$ is the average of measured yield.

3. Results

3.1. Relationships Between Yield and Vegetation Indices (VIs) and Textural Indices (TIs)

Figure 5 illustrates the distribution of the coefficient of determination (R^2) obtained from linear regressions between rice yield and various VIs across different phenological stages during the 2022 and 2023 growing seasons, as well as for the combined analysis of both years. The violin plots are color-coded by season (green = 2022, orange = 2023, blue = 2022–2023), which clarifies the temporal origin of each distribution.

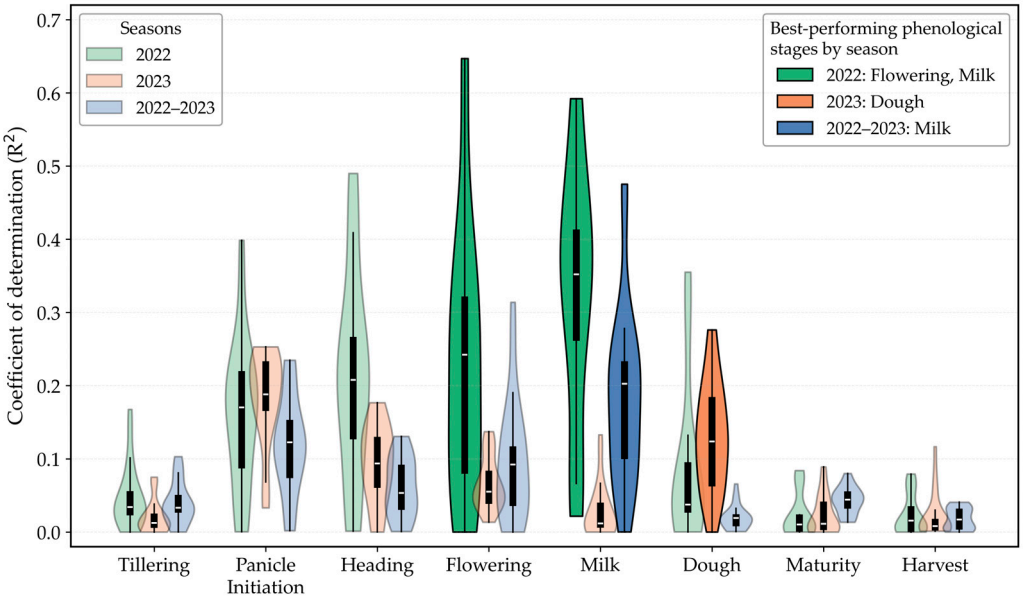


Figure 5. Phenological analysis of the coefficient of determination (R^2) between vegetation indices (VIs) and rice yield. The violin plots show the distribution of R^2 values across phenological stages for the 2022 season (green), the 2023 season (orange), and the combined 2022–2023 dataset (blue). Intermediate stages (Flowering, Milk, and Dough) yielded the highest predictive performance, while early and late stages exhibited consistently lower explanatory power.

In 2022, the Flowering and Milk stages achieved the highest R^2 values, highlighting their suitability for calibrating yield prediction models. In 2023, the Dough stage showed outstanding performance, while the combined 2022–2023 analysis confirmed the Milk stage as the most consistent predictor.

Early stages such as Tillering, Panicle Initiation, and Heading exhibited consistently lower R^2 values in all scenarios, indicating limited predictive capability. The variability observed between seasons and stages reflects the influence of climatic conditions and crop management practices on the relationship between VIs and yield.

These results emphasize that intermediate phenological stages (Flowering, Milk, and Dough) are critical for achieving accurate and reliable yield predictions. Selecting these stages for model calibration and validation is therefore essential to improve prediction performance at the plot scale.

3.2. Rice Yield–VI Correlations Across Phenological Stages

Figure 6 presents Pearson correlation matrices between rice yield and various vegetation indices (VIs) for four key phenological scenarios identified in the previous analysis: Flowering (2022), Milk (2022), Dough (2023), and the combined analysis of Milk (2022–2023). Correlations were calculated only with the plots available for each stage and season,

and evaluated under different significance levels (ns, $p \geq 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

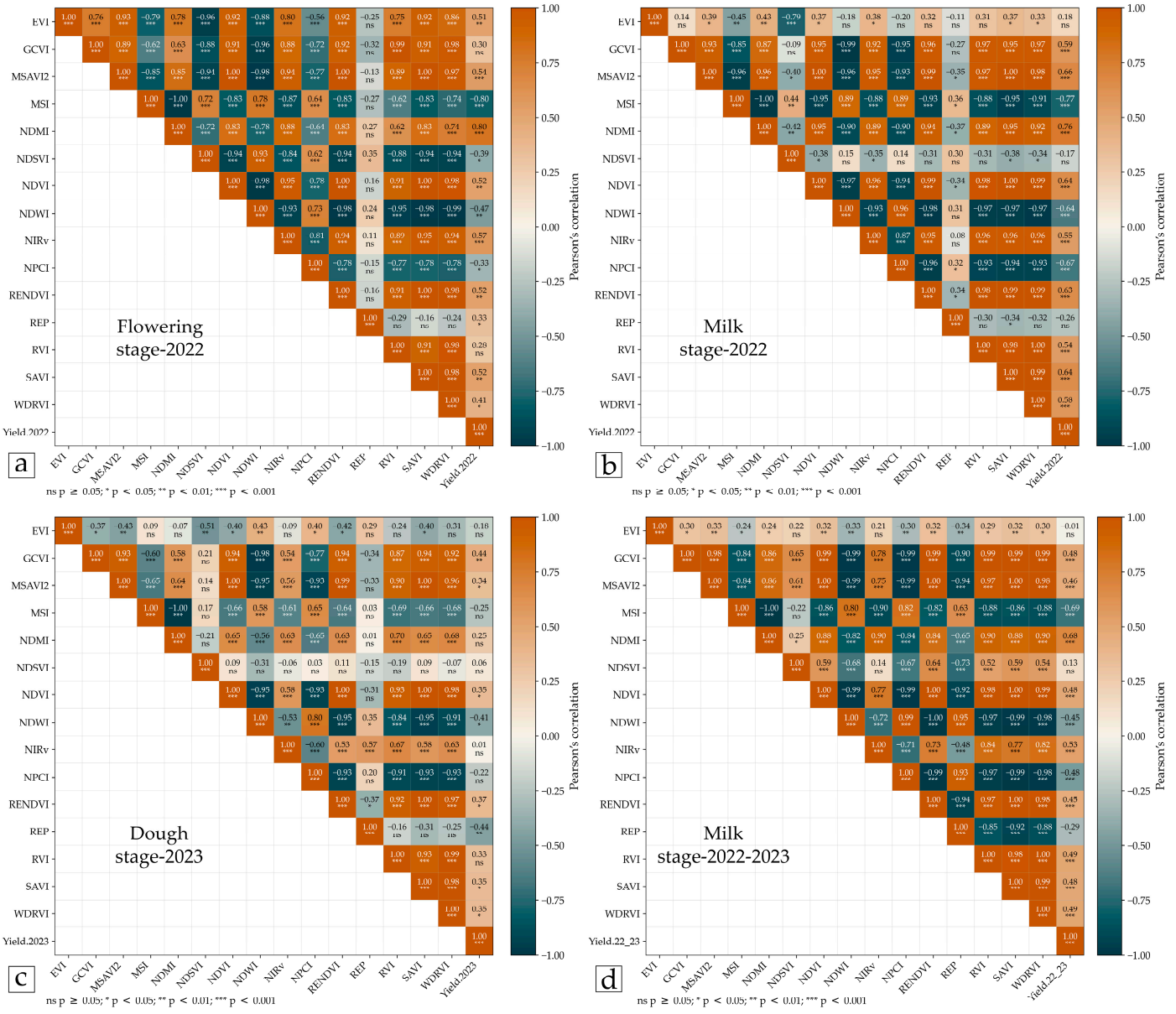


Figure 6. Pearson correlation between rice yield and different vegetation indices (VIs) at key phenological stages: (**a**) Flowering stage—2022, (**b**) Milk stage—2022, (**c**) Dough stage—2023, and (**d**) Milk stage—combined analysis for 2022–2023.

In Flowering (2022) (Figure 6a), a group of indices showed strong and highly significant correlations ($p < 0.001$) with yield, notably NDMI (r = 0.80 ***) and MSI (r = −0.80 ***), both associated with reflectance in the SWIR region and thus sensitive to canopy water content. MSAVI2 (r = 0.54 ***) also showed a significant positive correlation, whereas GCVI had a lower, non-significant value (r = 0.30 ns), suggesting that overall vegetative vigor is less determinant in this stage compared to water-related variables.

In Milk (2022) (Figure 6b), the association pattern remained consistent, with NDMI (r = 0.76 ***) and MSI (r = −0.77 ***) as the most yield-related indices, followed by MSAVI2 (r = 0.66 ***) and GCVI (r = 0.59 ***). This reinforces the idea that, during the grain-filling stage, indices combining near-infrared and SWIR information more accurately capture crop water and structural status.

In Dough (2023) (Figure 6c), correlations were notably weaker compared to 2022. The highest value was for GCVI (r = 0.44 **), while MSI showed a negative but non-significant value (r = −0.25 ns). This overall reduction may be linked to specific meteorological and management conditions during the 2023 season that attenuated spectral sensitivity at this advanced stage.

Finally, in the combined Milk analysis (2022–2023) (Figure 6d), the interannual stability of water-related indices was confirmed: MSI (r = −0.69 ***) and NDMI (r = 0.68 ***), maintaining high and consistent correlations with yield. This supports their utility as robust indicators for multi-temporal predictive models, even under climate variability between seasons.

Overall, these results—supported by detailed values in Table S1—reinforce the hypothesis that indices related to canopy water content and structure (e.g., NDMI and MSI) have the strongest and most stable relationship with rice yield, particularly in intermediate phenological stages. In contrast, indices such as MSAVI2 and GCVI provide complementary value, especially in scenarios where vegetative vigor plays a greater role. This pattern suggests that combining water- and vigor-based metrics could optimize rice yield prediction models under contrasting management and climatic conditions.

Although correlation analyses were conducted across all phenological stages and both seasons, Figure 6 highlights only the stage–year combinations with the highest predictive ability (Flowering 2022, Milk 2022, Dough 2023, and Milk 2022–2023). Other combinations, such as Flowering 2023 or Dough 2022, were not included in the main figure because they were strongly affected by cloud contamination and yielded substantially weaker or non-significant correlations. Full correlation results for all stage–year combinations are provided in Table S1. This approach ensures transparency while focusing on the most relevant scenarios for rice yield modeling.

3.3. Performance of Machine Learning Models for Rice Yield Prediction in 2022, 2023 and Their Combination

3.3.1. Prediction Models Using Multiple Linear Regression (MLR) and Support Vector Machines (SVR) with Sequential Forward Selection (SFS)

At this stage, predictive models were developed using MLR and SVR with both linear and radial basis function (RBF) kernels, incorporating the Sequential Forward Selection (SFS) technique. This approach progressively identified the spectral variable combinations that minimized the root mean square error (RMSE) through leave-one-out cross-validation (LOOCV).

Figure 7 presents the best-performing configurations for MLR, linear SVR, and RBF SVR, illustrating the RMSE evolution as a function of the number of selected variables and highlighting the combinations that defined the "best model" in each case. The in-plot labels specify the spectral indices contributing to the optimal model performance, with results shown for the Milk and Flowering stages. Overall, these findings indicate that a small number of carefully selected variables can achieve competitive predictive accuracy, avoiding feature overload and enhancing the model's generalization capability.

Complete results for other model–year–phenological stage combinations are provided in the Supplementary Material (Figures S1–S4).

The scatter plots illustrate the relationship between observed and predicted rice yields for the best-performing MLR, linear SVR, and SVR-rbf models in the 2022 season, selected through Sequential Forward Selection (SFS). In the Milk stage, the MLR model incorporated MSAVI2, NDMI, NDVI, and NDWI, achieving an R^2_CV of 0.68, an RMSE_CV of 1.30 t ha^{-1}, and an rRMSE_CV of 12.64%. In the same stage, the linear SVR model selected MSI, NDSVI, NDWI, and RVI, reaching an R^2_CV of 0.67, an RMSE_CV of 1.33 t ha^{-1}, and an rRMSE_CV of 12.94%. For the Flowering stage, the RBF SVR model used EVI,

NDMI, NDSVI, and NPCI, obtaining an R^2_CV of 0.66, an RMSE_CV of 1.34 t ha^{-1}, and an rRMSE_CV of 13.05%. Across all cases, the points closely followed the 1:1 reference line, indicating strong agreement between observed and predicted yields. These results highlight the models' ability to predict rice yield accurately using a small set of key spectral variables (Figure 8).

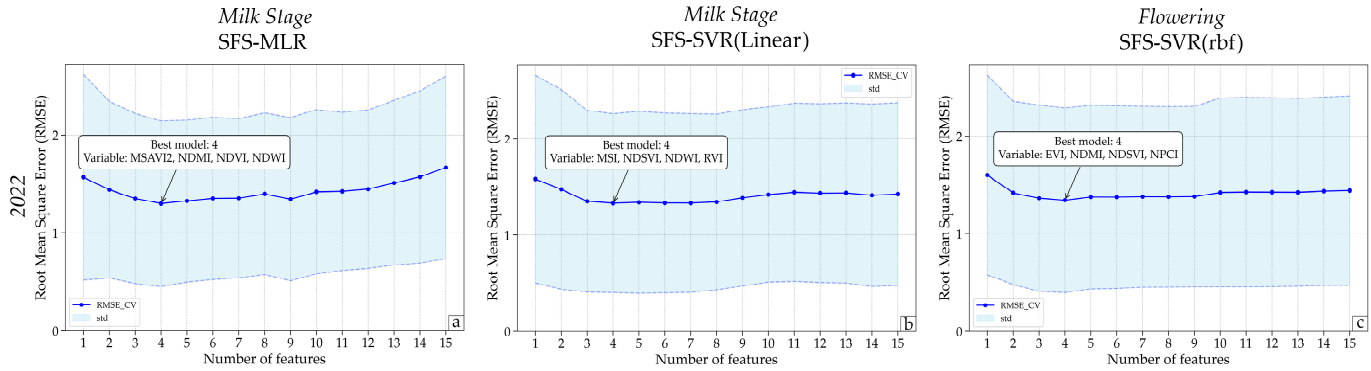


Figure 7. Evolution of root mean square error (RMSE) as a function of the number of features selected using Sequential Forward Selection (SFS) for (**a**) MLR in the Milk stage—2022, (**b**) SVR with linear kernel in the Milk stage—2022, and (**c**) SVR with RBF kernel in the Flowering stage—2022. The best model in each case is indicated, along with the corresponding selected variables.

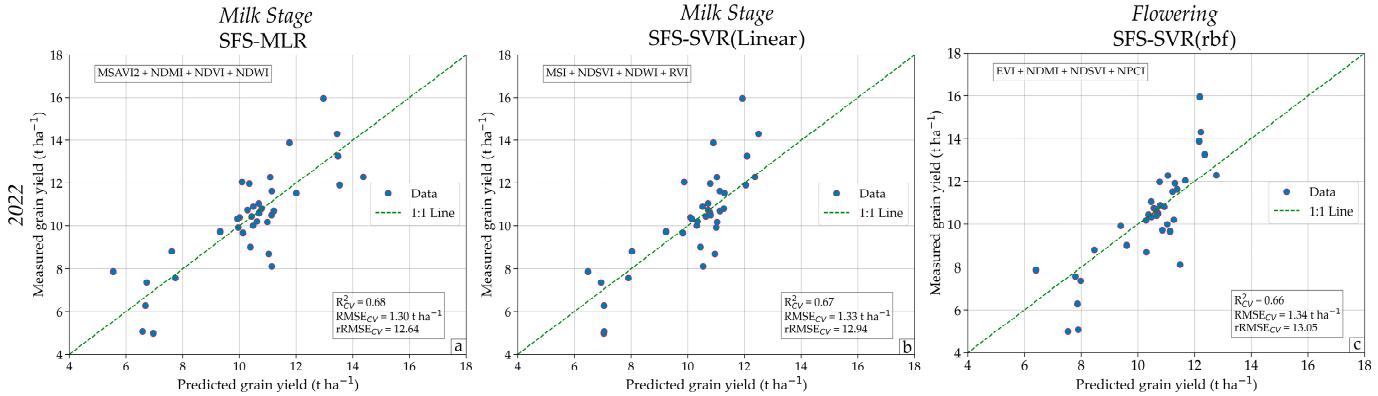


Figure 8. Relationship between observed and predicted rice yield for the best-performing models in the 2022 season, selected using Sequential Forward Selection (SFS): (**a**) Multiple Linear Regression (MLR) in the Milk stage; (**b**) linear Support Vector Regression (SVR) in the Milk stage; and (**c**) radial basis function (RBF) SVR in the Flowering stage.

In 2022, the MLR, linear SVR, and RBF SVR models achieved the highest R^2_CV values and the lowest RMSE_CV across the entire study period. In 2023, model performance was comparatively lower, and for the combined 2022–2023 dataset, predictive accuracy was slightly reduced compared to the 2022 season.

3.3.2. Performance of PLSR: Cross-Validation Results and SHAP-Based Importance

In the 2022 Flowering stage, the PLSR model achieved its optimal performance using six latent components, striking a balance between predictive capability and overfitting control (Figure 9a,b). The evolution of R^2 and RMSE indicates that a reduced number of components enhances the model's generalization ability.

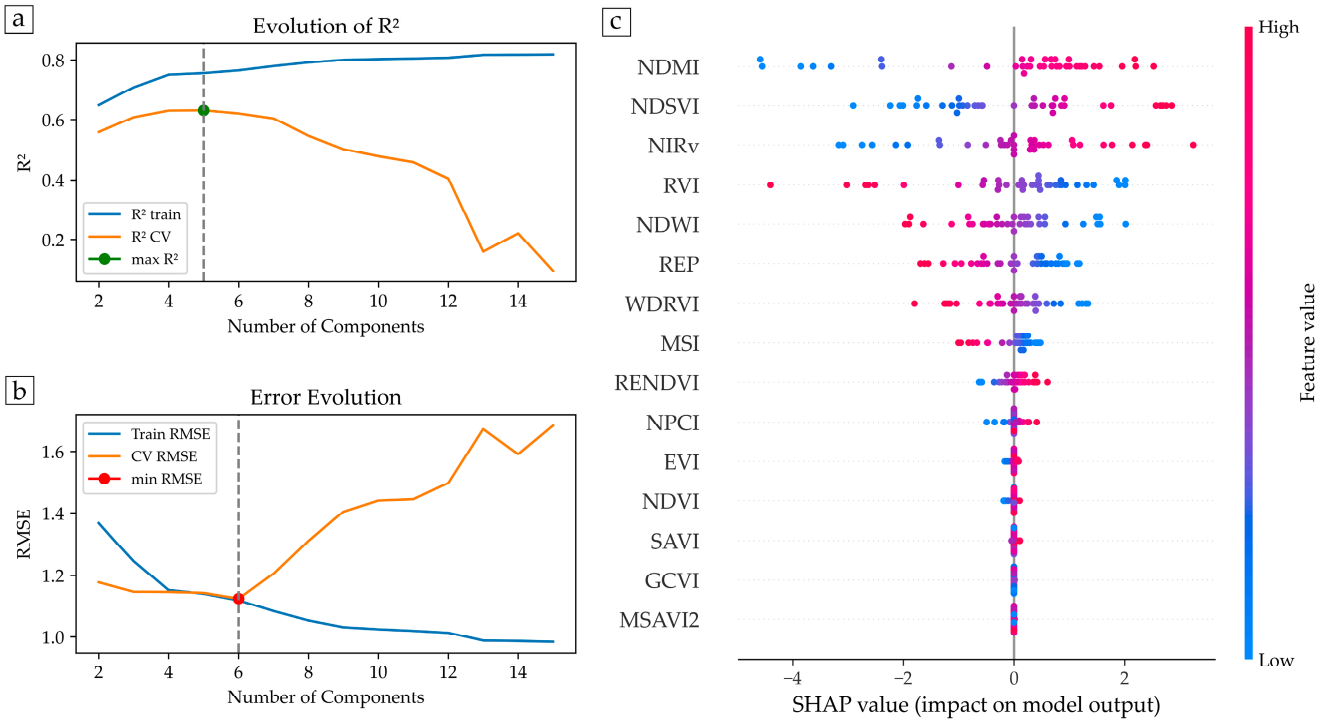


Figure 9. Performance of the PLSR model and variable importance for rice yield prediction at the Flowering stage (2022): (**a**) evolution of R^2 with the number of latent components; (**b**) evolution of RMSE with the number of latent components; and (**c**) SHAP-based feature importance indicating the contribution of vegetation indices to the model's predictive output.

The SHAP-based feature importance analysis (Figure 9c) revealed that water-related indices (NDMI, NDSVI, NDWI, and MSI) and vegetation vigor metrics (NIRv, RVI, MSAVI2) were the primary determinants in yield prediction. This finding underscores the importance of integrating complementary spectral metrics to improve plot-level rice yield estimation.

At the Flowering stage of 2022, the PLSR model exhibited a strong agreement between predicted and field-measured grain yield, with data points closely aligned along the 1:1 reference line (Figure 10). The model achieved an R^2_CV of 0.68, an RMSE_CV of 1.31 t ha^{-1}, and an rRMSE_CV of 12.74%, indicating good predictive performance and a moderate relative error. These results confirm the suitability of PLSR for estimating rice yield from selected spectral indices at this phenological stage.

Overall, the PLSR results for the flowering stage (2022) indicate that using a moderate number of latent components maximizes the model's generalization ability while avoiding overfitting and maintaining an acceptable relative error. The prominent importance of indices related to canopy water content (NDMI, NDSVI, NDWI, MSI) and vegetative vigor (NIRv, RVI, MSAVI2) underscores the value of integrating complementary spectral metrics to more accurately capture the physiological and structural status of the crop. These findings are consistent with previous studies reporting the higher predictive power of water-sensitive indices during critical reproductive stages, thereby reinforcing the potential of PLSR as a robust tool for rice yield prediction using Sentinel-2 data.

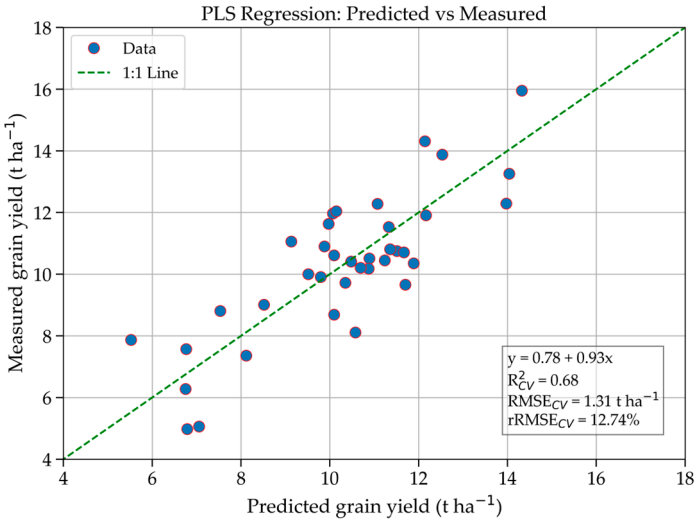


Figure 10. Relationship between predicted and measured rice grain yield using the PLSR model at the flowering stage (2022), showing the 1:1 reference line and performance metrics R^2_CV, RMSE_CV, and rRMSE_CV.

3.3.3. Ensemble Learning Models: Random Forest (RF) and Extreme Gradient Boosting (XGBoost)

At the flowering stage in 2022, the Random Forest (RF) model achieved its optimal configuration with 16 trees and a single spectral variable, as determined through grid search optimization (Table S3). The evolution of out-of-bag (OOB) error and RMSE (Figure 11a,b) shows that increasing the number of trees rapidly reduces the initial error until stabilization, while adding more variables beyond the optimal configuration does not improve performance. This suggests that a small, well-selected set of predictors is sufficient to maximize the model's generalization ability.

The variable importance analysis assessed using %IncMSE and IncNodePurity (Figure 11c), identified NDMI as the most influential index for yield prediction, followed by NDSVI, WDRVI, RVI, and NDVI. This finding confirms the relevance of water-content and vegetation-vigor indices previously highlighted in the correlation analyses. The same pattern was observed across other phenological stages (Tables S4 and S5), although the optimal number of trees and variables varied, highlighting the need to adjust model parameters to each specific crop growth phase.

In the flowering stage of 2022, the Random Forest (RF) model exhibited a moderate agreement between predicted and measured grain yield, with most data points distributed near the 1:1 reference line (Figure 12). The model achieved an R^2_CV of 0.57, an RMSE_CV of 1.52 t ha^{-1}, and an rRMSE_CV of 14.73%. While these results indicate an acceptable predictive capacity, they were lower than those obtained with PLSR for the same phenological stage. This suggests that the optimal RF configuration—based on a reduced number of trees and spectral variables—may be sensitive to inter-plot variability and could benefit from incorporating a larger set of predictors to improve accuracy under contrasting management and climatic conditions.

Overall, the RF results at the flowering stage (2022) indicate that using a reduced number of trees (16) and a single spectral variable optimized model generalization, albeit with a moderate predictive capacity compared to PLSR. The identified importance of NDMI, NDSVI, WDRVI, RVI, and NDVI confirms the predominant role of water-sensitive and vegetative vigor indices in yield estimation, consistent with the outcomes of correlation and PLSR analyses. However, the lower accuracy achieved during validation suggests that maximizing the potential of RF under variable management and climatic conditions may

require expanding the predictor set or refining hyperparameters beyond the tested grid search configuration.

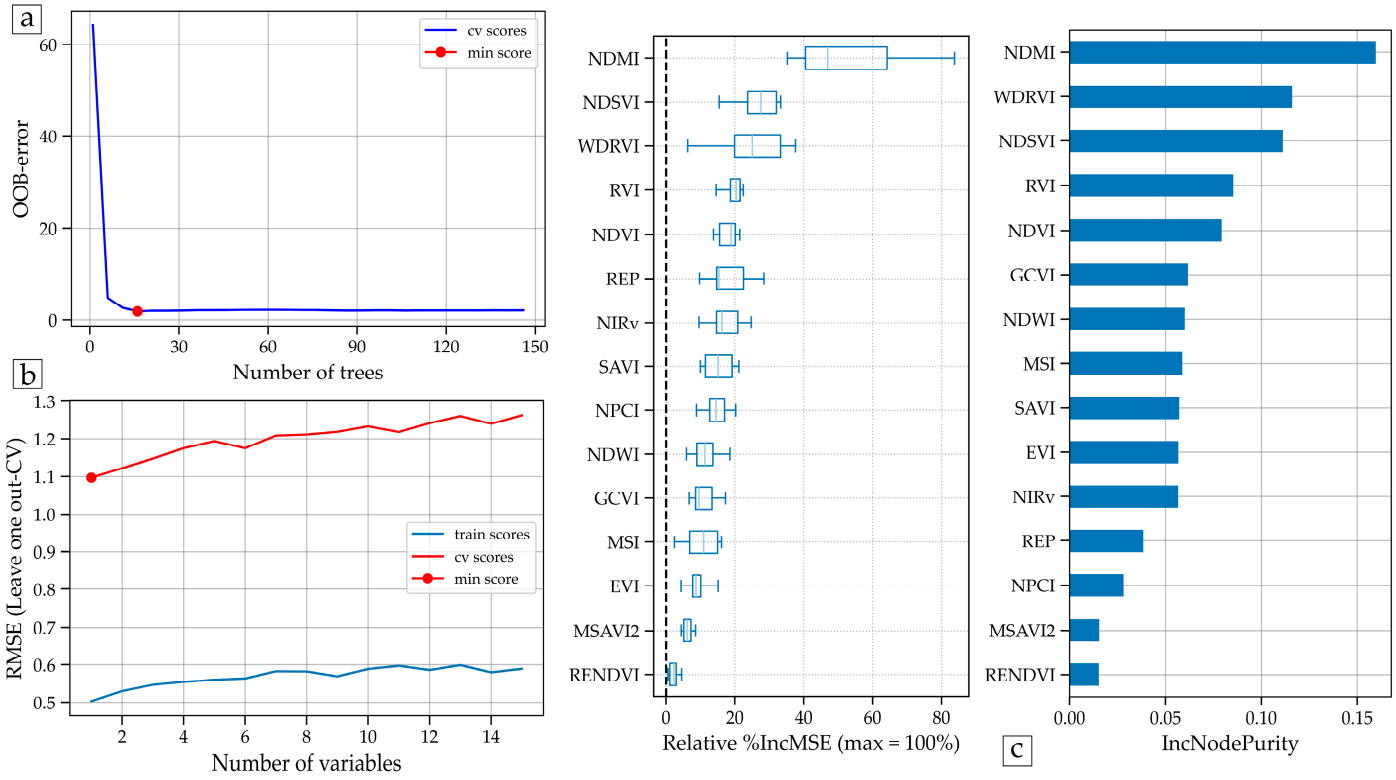


Figure 11. Performance and variable importance of the Random Forest (RF) model for rice yield prediction at the flowering stage (2022): (**a**) evolution of out-of-bag (OOB) error with the number of trees; (**b**) evolution of RMSE with the number of variables; and (**c**) variable importance measured by relative %IncMSE (boxplots, scaled to the maximum predictor = 100%) and IncNodePurity (bar chart, unitless), highlighting the contribution of vegetation indices to model performance.

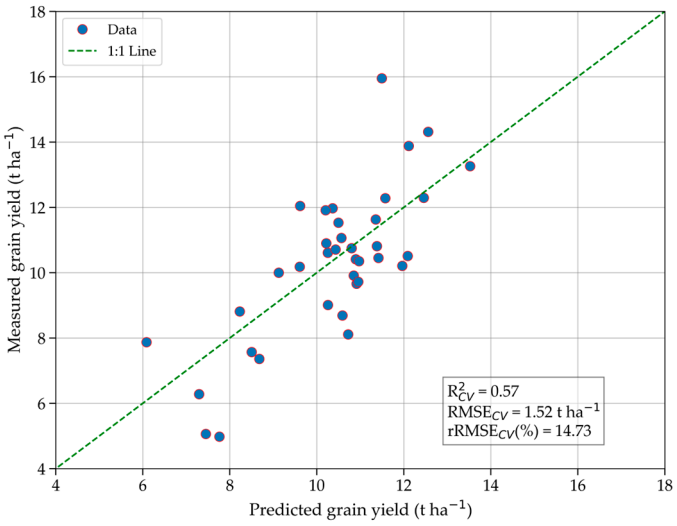


Figure 12. Relationship between predicted and measured rice grain yield using the Random Forest (RF) model at the flowering stage (2022), showing the 1:1 reference line and performance metrics R^2_CV, RMSE_CV, and rRMSE_CV.

In the milk stage of 2022, the Extreme Gradient Boosting (XGBoost) model exhibited a rapid decrease in cross-validation error (RMSE) as the number of trees increased, stabilizing

at approximately 100 iterations (Figure 13a). The assessment of error as a function of maximum tree depth indicated that the optimal value was achieved with a relatively shallow depth, thereby preventing overfitting and enhancing model generalization (Figure 13b). The permutation importance analysis (Figure 13c) revealed that EVI, MSAVI2, and MSI were the most influential predictors in the model, followed by GCVI, REP, NDSVI, and NIRv. The predominance of vegetation vigor metrics (EVI, MSAVI2, GCVI) alongside canopy water-content and structural indicators (MSI, NDSVI) underscores the importance of integrating complementary spectral information to accurately capture the physiological condition of the crop during this critical grain-filling stage.

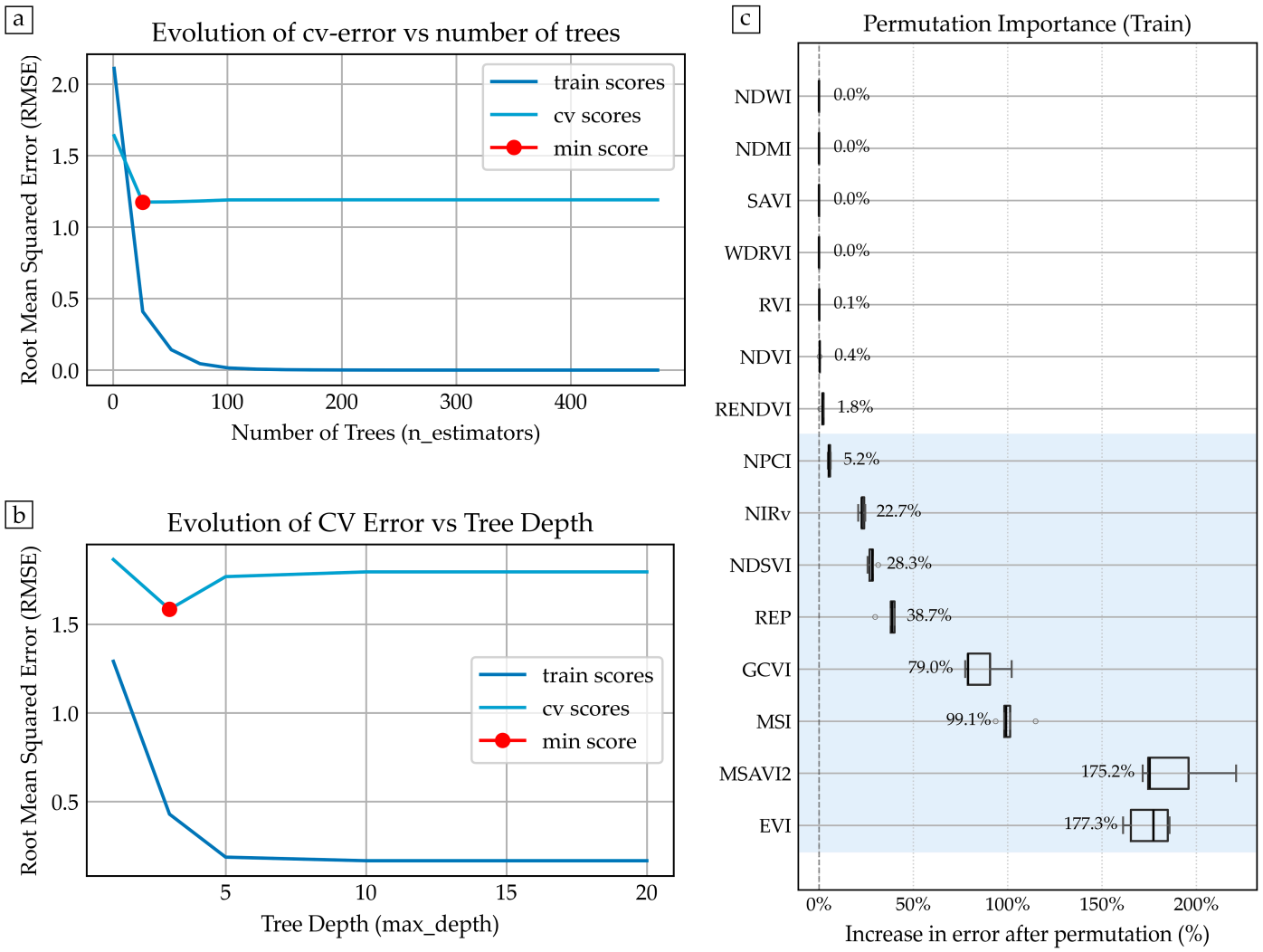


Figure 13. XGBoost model optimization for rice yield prediction at the milk stage (2022): (**a**) evolution of CV error vs. number of trees; (**b**) evolution of CV error vs. tree depth; and (**c**) permutation-based variable importance.

In the milk stage of 2022, the XGBoost model exhibited a moderate agreement between predicted and measured grain yield, with most data points distributed near the 1:1 reference line (Figure 14). The model achieved an R^2_CV of 0.55, an RMSE_CV of 1.55 t t ha^{-1}, and an rRMSE_CV of 15.08%. While these results indicate an acceptable predictive capacity, they were lower than those obtained with PLSR at the flowering stage and comparable to RF in its optimal configuration. This suggests that the performance of XGBoost during the grain-filling phase could be enhanced through more refined hyperparameter tuning and

the inclusion of additional variables that better capture intra-plot variability and canopy water status under diverse management and climatic conditions.

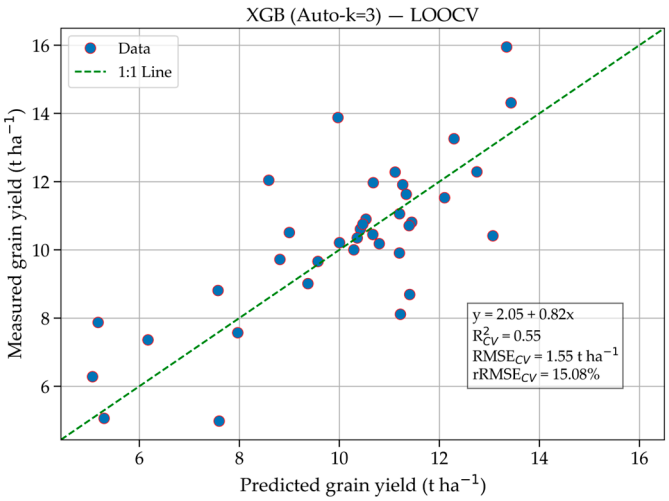


Figure 14. Relationship between predicted and measured rice grain yield using the XGBoost model at the milk stage (2022), showing the 1:1 reference line and performance metrics R^2_CV, RMSE_CV, and rRMSE_CV.

Overall, the XGBoost results at the milk stage (2022) show that optimizing the number of trees and maximum depth substantially reduced cross-validation error, achieving a balance between predictive accuracy and overfitting control. The high importance assigned to EVI, MSAVI2, and MSI highlights the crucial role of vegetative vigor and canopy water content metrics during the grain-filling phase, consistent with patterns observed for PLSR and RF. Nonetheless, the moderate predictive capacity suggests that incorporating a broader set of predictors or applying more refined hyperparameter tuning could enhance model performance under diverse management and climatic scenarios.

3.3.4. Performance of the Yield Prediction Models

Table 4 summarizes the predictive performance of the machine learning models evaluated using leave-one-out cross-validation (LOOCV) for the 2022 and 2023 growing seasons, as well as for the combined 2022–2023 dataset. These results complement the training metrics presented in Table S2.

In 2022, the highest performance at the flowering stage was achieved by PLSR (R^2_CV = 0.68, RMSE_CV = 1.31 t ha^{-1}, rRMSE_CV = 12.74%), slightly outperforming MLR (R^2_CV = 0.67) and SVR-rbf (R^2_CV = 0.66). At the milk stage (2022), MLR obtained the best fit (R^2_CV = 0.68; RMSE_CV = 1.30 t ha^{-1}), closely followed by SVR-linear (R^2_CV = 0.67).

In 2023, at the dough stage, the best-performing model was MLR (R^2_CV = 0.24; RMSE_CV = 0.88 t ha^{-1}). However, all algorithms exhibited relatively low R^2 values, indicating a reduced predictive capacity for this phenological phase and growing season.

For the combined analysis at the milk stage (2022–2023), MLR again achieved the highest performance (R^2_CV = 0.63; RMSE_CV = 1.15 t ha^{-1}), with moderate differences compared to SVR-linear (R^2_CV = 0.58) and SVR-rbf (R^2_CV = 0.57).

Table 4. Predictive performance of machine learning models for rice yield estimation across phenological stages (Flowering, Milk, Dough) during 2022, 2023, and combined (2022–2023) seasons using leave-one-out cross-validation (LOOCV).

Machine Learning Models	Validation (Leave One out CV)											
	2022						2023			2022–2023		
	Flowering			Milk			Dough			Milk		
	R^2_CV	RMSE_CV (t ha^{-1})	rRMSE_CV (%)	R^2_CV	RMSE_CV (t ha^{-1})	rRMSE_CV (%)	R^2_CV	RMSE_CV (t ha^{-1})	rRMSE_CV (%)	R^2_CV	RMSE_CV (t ha^{-1})	rRMSE_CV (%)
Vegetation Index (VI) Multiple Linear Regression (MLR)	0.67	1.33 (0.79)	12.94 (7.62)	0.68	1.30 (0.85)	12.64 (8.23)	0.24	0.88 (0.56)	9.56 (6.10)	0.63	1.15 (0.75)	11.73 (7.69)
Support Vector Machine (SVR-linear)	0.57	1.51 (0.82)	14.70 (7.95)	0.67	1.33 (0.93)	12.94 (9.03)	0.13	0.95 (0.62)	10.26 (6.73)	0.58	1.21 (0.88)	12.41 (9.04)
Support Vector Machine (SVR-rbf)	0.66	1.34 (0.95)	13.05 (9.21)	0.63	1.41 (1.09)	13.70 (10.60)	0.10	0.97 (0.60)	10.46 (6.52)	0.57	1.23 (0.91)	12.61 (9.35)
Partial Least Squares Regression (PLSR)	0.68	1.31 (0.69)	12.74 (6.71)	0.56	1.53 (0.97)	14.88 (9.39)	0.14	0.94 (0.59)	10.19 (6.39)	0.49	1.34 (0.94)	13.69 (9.57)
Random Forest (RF)	0.57	1.52 (0.89)	14.73 (8.66)	0.54	1.56 (1.09)	15.18 (10.54)	−0.19 *	1.11 (0.69)	12.01 (7.46)	0.46	1.38 (0.97)	14.13 (9.91)
Extreme Gradient Boosting (XGBoost)	0.51	1.61 (1.00)	15.63 (9.72)	0.55	1.55 (1.06)	15.08 (10.26)	−0.27 *	1.14 (0.68)	12.39 (7.32)	0.43	1.42 (0.92)	14.48 (9.36)

Note: R^2_CV values were computed as 1—(SSE/TSS) under LOOCV. * Negative values indicate cases where prediction errors exceeded the variance of the observed yields, meaning the model performed worse than the mean baseline.

When comparing these validation results with the training metrics in Table S2, it is evident that XGBoost and RF, despite achieving very high R^2 values during training (up to 1.00 for XGBoost and 0.94 for RF), showed a substantial drop in performance during validation. This pattern suggests potential overfitting, likely linked to the high model complexity and the limited sample size. In contrast, MLR and PLSR maintained more consistent results between training and validation, indicating greater generalization capacity under the LOOCV scheme.

In the 2023 season, all models yielded lower R^2 values compared to 2022, particularly at the dough stage. This decrease may be associated with interannual variability in climatic and management conditions, as well as with reduced availability or quality of satellite observations during critical crop growth stages. Such differences may have weakened the relationship between spectral indices and yield, thereby reducing the models' predictive power.

3.3.5. Spatial Prediction of Rice Yield at Plot Scale Using Sentinel-2 Data

The MLR model calibrated at the milk stage (2022) showed the best performance (R^2 = 0.77; RMSE = 1.11 t ha^{-1}) and was therefore selected for spatial prediction. Figure 15a depicts the predicted rice yield at the plot scale, with values ranging from 6.0 to 14.0 t ha^{-1} and a clear differentiation among fields. The residual map (Figure 15b) indicates deviations between observed and predicted values ranging from −2.6 to +2.6 t ha^{-1}. To further characterize error distribution, a histogram of residuals (Figure 16) was included, showing that most residuals clustered around zero and the majority were contained within ±0.5 t ha^{-1}. These results confirm the reliability of the MLR model at the milk stage for generating accurate yield maps and demonstrate the potential of Sentinel-2 data to support precision rice management in irrigated systems.

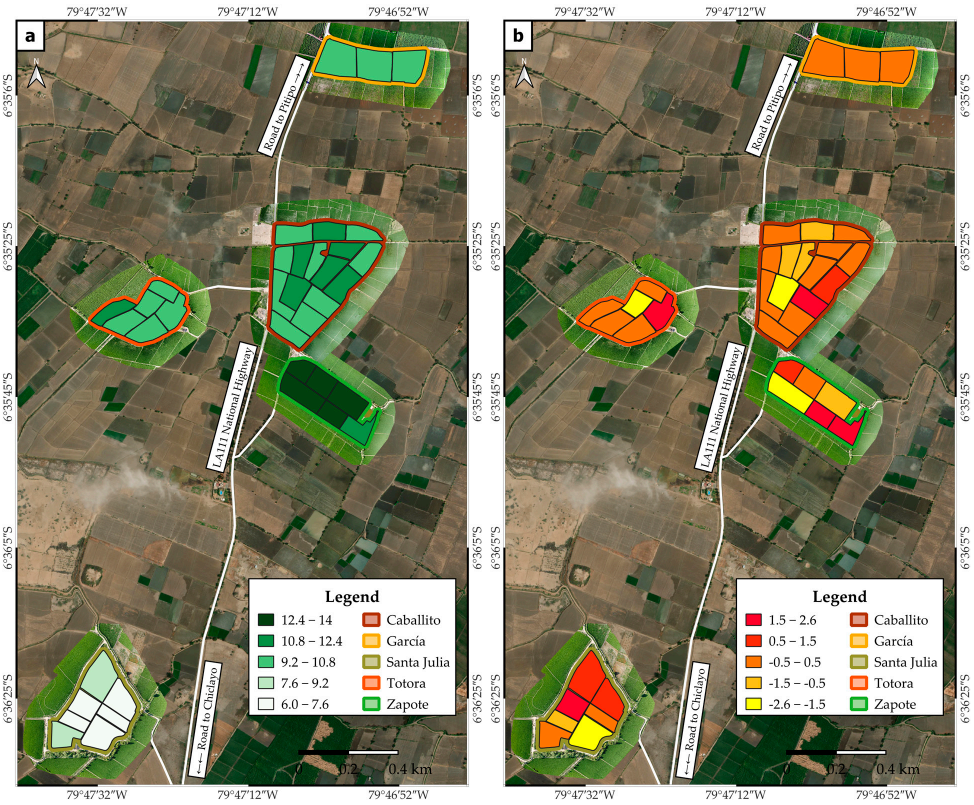


Figure 15. Spatial prediction of rice yield at plot scale using Sentinel-2 data with the best-performing model (MLR, milk stage 2022). (**a**) Predicted rice yield (t ha^{-1}). (**b**) Residuals (observed–predicted, t ha^{-1}).

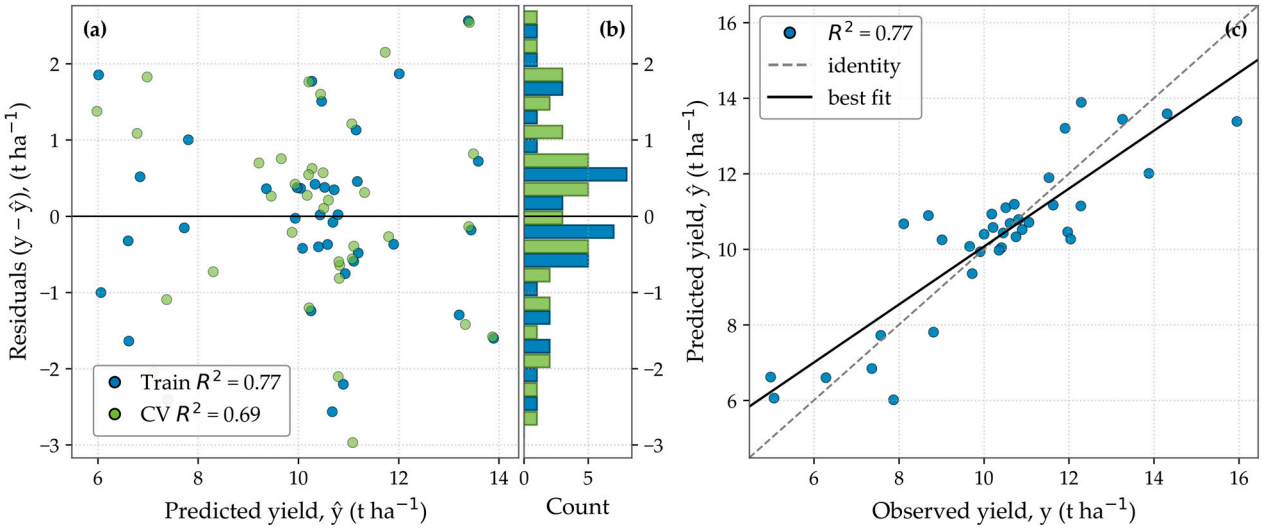


Figure 16. Residual analysis and model performance for the best-performing model (MLR, milk stage 2022). (**a**) Scatterplot of residuals (observed–predicted, t ha^{-1}) against predicted yield, showing random distribution without systematic bias. (**b**) Histogram of residuals, where blue bars represent the training set and green bars represent the cross-validation (CV) set, highlighting that most errors clustered within ±0.5 t ha^{-1} despite a few extreme values reaching ±2.6 t ha^{-1}. (**c**) Observed versus predicted rice yield with identity (dashed) and best-fit (solid) lines, confirming strong agreement (R^2 = 0.77).

4. Discussion

4.1. Phenological Stages Govern the Strength and Stability of Yield–VI Relationships

Our findings confirm that rice yield prediction accuracy is strongly influenced by phenological stage, with flowering and milk consistently outperforming early vegetative and late reproductive phases. In 2022, the highest LOOCV performance was achieved by PLSR at flowering (R^2_CV = 0.68; RMSE_CV = 1.31 t ha^{-1}; rRMSE_CV = 12.74%) and MLR at milk (R^2_CV = 0.68; RMSE_CV = 1.30 t ha^{-1}; rRMSE_CV = 12.64%). This aligns with previous Sentinel-2 studies showing that reproductive stages—particularly flowering—provide optimal predictive power due to stable canopy structure and high sensitivity to water status [9,61,62]. By contrast, dough-stage predictions in 2023 showed markedly lower accuracy (e.g., MLR: R^2_CV = 0.24; RMSE_CV = 0.88 t ha^{-1}), reflecting both phenological and climatic effects. The spatial distribution of rice yield predicted with the best-performing model (MLR at milk 2022) is shown in Figure 15, confirming the model's ability to capture intra-plot yield variability and residual patterns.

4.2. Phenological Sensitivity and Key Spectral Predictors

Spectral indices linked to canopy water content and SWIR reflectance—particularly NDMI and MSI—consistently emerged as the top predictors across high-performing models, followed by vigor-related metrics such as MSAVI2 and GCVI. In 2022, NDMI and MSI reached correlations of r = 0.80 and r = −0.80 with yield at flowering, and r = 0.76 and r = −0.77 at milk. These results reinforce the importance of combining physiological (water content) and structural (vigor) cues, echoing findings in multi-temporal rice yield modeling [5,12,24]. The stability of NDMI and MSI in the combined 2022–2023 milk-stage analysis (r = 0.68 and r = −0.69) supports their robustness under interannual variability.

4.3. Model Behavior: Parsimony vs. Complexity

While ensemble models such as RF and XGBoost achieved high in-sample accuracy (e.g., RF at flowering 2022: R^2_train > 0.90), they suffered validation performance drops

under LOOCV (RF: R^2_CV = 0.57; RMSE_CV = 1.52 t ha^{-1}), indicative of overfitting in high-dimensional, small-sample contexts. Conversely, parsimonious models like MLR and PLSR maintained stable validation accuracy, in line with documented benefits of predictor selection in remote sensing yield modeling [14,51,63]. Our use of sequential forward selection (SFS) mitigated redundancy, yielding compact predictor sets (3–5 VIs) without compromising accuracy.

4.4. Interannual Variability and Climatic Context

Marked performance differences between 2022 and 2023 underscore the role of climatic variability in rice yield prediction. In particular, Dough-stage models in 2023 showed very low accuracy (best R^2_CV = 0.14), coinciding with higher average maximum temperatures and reduced relative humidity during grain filling. Comparable seasonal sensitivity has been widely reported in rice yield–VI relationships, where environmental stress reduces the strength of spectral–yield coupling [1,8,64,65].

Although we also tested models integrating meteorological variables from the INIA–Vista Florida station and reanalysis products (ERA5, PISCO), their coarse spatial resolution provided only uniform values across all commercial plots. As a result, they did not improve predictive performance, underscoring that interannual variability could not be fully captured without finer-scale climate information.

Integrating ancillary climate variables or management indicators therefore remains a promising pathway to enhance model robustness across years [66]. For example, Quille-Mamani et al. [65] demonstrated that models excluding climate covariates systematically underperformed under stress conditions and emphasized the need to incorporate them in operational monitoring frameworks. Similarly, Guzman-Lopez et al. [67] conducted a multi-regional analysis of Peruvian rice systems and showed that precipitation and temperature dynamics—including velocity and acceleration terms—significantly improved model accuracy and revealed causal climate–yield linkages in the Granger sense. Collectively, these results highlight that the integration of climate variables with spectral indices is essential to strengthen prediction frameworks and ensure stability of rice yield models in coastal Peru.

4.5. Positioning Within Prior Work and Implications

This study reinforces the evidence that Sentinel-2 VIs—when selected for phenological relevance and modeled with appropriate regularization—can match or surpass complex learners in yield prediction [5,12,60]. Our fully reproducible GEE–Python pipeline integrates spatial harmonization, SFS, and LOOCV, aligning with best practices for operational remote sensing-based yield forecasting [17]. The findings support the feasibility of deploying such workflows in water-limited rice systems without UAV data dependency.

4.6. Limitations and Future Work

The main limitations of this study include the restricted sample size per phenological stage, occasional gaps in Sentinel-2 observations due to cloud cover, and the sensitivity of models to intra-plot variability. Moreover, although we tested models incorporating meteorological information from the INIA–Vista Florida station, as well as reanalysis and gridded products such as ERA5 and PISCO, these datasets provided only spatially aggregated values that remained constant across all commercial plots. Consequently, their integration did not enhance predictive performance, as the absence of intra-plot variability limited their explanatory contribution. This outcome reflects a broader structural challenge in Peru, where the sparse distribution of operational meteorological stations and the coarse resolution of available reanalysis products hinder the effective integration of climatic covariates into plot-scale crop models.

Future work should therefore prioritize (i) expanding the temporal and spatial coverage of ground-truth data to improve calibration and validation, (ii) adopting multi-sensor fusion approaches (e.g., Sentinel-1 SAR, UAV-based thermal and multispectral imagery), (iii) integrating texture- and phenology-based features to strengthen spatial transferability, and (iv) deploying higher-resolution meteorological monitoring systems, ideally with distributed sensors at the plot or irrigation-command scale, to capture microclimatic variability. These strategies align with recent recommendations for enhancing the robustness and scalability of yield prediction frameworks in heterogeneous agroecosystems [5,66,68]. By combining fine-scale climatic observations with spectral indicators, future modeling efforts will be better positioned to disentangle the effects of environmental stressors and to improve the stability and reliability of rice yield predictions across contrasting seasons and management conditions.

5. Conclusions

This study demonstrates that Sentinel-2 vegetation indices (VIs) can be used to predict rice yield at the plot scale with competitive accuracy when model calibration is performed at phenological stages that capture key physiological and structural transitions, particularly flowering and milk. Across the two evaluated growing seasons (2022 and 2023) in the Chancay–Lambayeque Valley, Peru, the best validation results were obtained with Partial Least Squares Regression (PLSR) at flowering in 2022 (R^2_CV = 0.68; RMSE_CV = 1.31 t ha^{-1}; rRMSE_CV = 12.74%) and Multiple Linear Regression (MLR) at milk (R^2_CV = 0.68; RMSE_CV = 1.30 t ha^{-1}; rRMSE_CV = 12.64%). These outcomes confirm that parsimonious models, supported by targeted variable selection, can match or even outperform more complex machine learning algorithms.

Spectral predictors associated with canopy water content (NDMI, MSI, NDWI, NDSVI) and vegetative vigor (EVI, MSAVI2, GCVI, NIRv) consistently emerged as the most relevant variables, as revealed by SHAP and permutation-based importance metrics. The stability of these predictors across phenological stages and seasons underscores their potential as robust indicators for operational yield monitoring. Conversely, ensemble learners such as Random Forest (RF) and Extreme Gradient Boosting (XGBoost) exhibited high training performance but notable drops in validation accuracy, reflecting overfitting risks when sample sizes are limited and predictor spaces are high-dimensional.

Interannual differences were evident, with lower predictive capacity in 2023—especially at the dough stage—likely driven by variations in climate, water management, and satellite observation conditions. This aligns with previous multi-season studies that have reported weakened yield–VI coupling under variable environmental conditions. Such variability highlights the need to incorporate complementary predictors, including meteorological and management data, to improve model transferability.

From a methodological perspective, the integration of Google Earth Engine for spectral data extraction, Sequential Forward Selection (SFS) for variable reduction, and Leave-One-Out Cross-Validation (LOOCV) for performance assessment offers a reproducible and computationally efficient framework for yield prediction in smallholder rice systems. These strategies are readily adaptable to other crops and regions where timely, low-cost yield estimates are critical for agricultural planning.

Future research should aim to (i) expand temporal coverage to multiple cropping cycles and climatic regimes, (ii) incorporate multi-sensor data fusion (e.g., Sentinel-1 SAR, UAV imagery), (iii) integrate texture- and phenology-based features, and (iv) evaluate spatial transferability to neighboring agroecological zones. By combining phenologically targeted spectral indicators with robust modeling strategies, operational yield forecasting

can be advanced to better support water resource management, food security, and climate resilience in irrigated rice systems.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/agriculture15192054/s1: Figure S1. Sequential forward feature selection (SFS) results for Multiple Linear Regression (MLR) and Support Vector Regression (SVR-linear, SVR-rbf) models using Sentinel-2 vegetation indices at the flowering stage (2022). The plots show RMSE variation under LOOCV as a function of the number of selected features, highlighting the best-performing model configuration in each case; Figure S2. Sequential forward feature selection (SFS) results for Multiple Linear Regression (MLR) and Support Vector Regression (SVR-linear, SVR-rbf) models using Sentinel-2 vegetation indices at the milk stage (2022). The plots show RMSE variation under LOOCV as a function of the number of selected features, highlighting the best-performing model configuration in each case; Figure S3. Sequential forward feature selection (SFS) results for Multiple Linear Regression (MLR) and Support Vector Regression (SVR-linear, SVR-rbf) models using Sentinel-2 vegetation indices at the dough stage (2023). The plots show RMSE variation under LOOCV as a function of the number of selected features, highlighting the best-performing model configuration in each case; Figure S4. Sequential forward feature selection (SFS) results for Multiple Linear Regression (MLR) and Support Vector Regression (SVR-linear, SVR-rbf) models using Sentinel-2 vegetation indices at the milk stage (2022–2023). The plots show RMSE variation under LOOCV as a function of the number of selected features, highlighting the best-performing model configuration in each case; Table S1. Pearson correlation coefficients (r) between rice yield and different vegetation indices (VIs) at key phenological stages—Flowering (2022), Milk (2022), Milk (2022–2023), and Dough (2023)—with indication of statistical significance levels (ns, *, **, *); Table S2. Predictive performance of machine learning models using Sentinel-2–derived vegetation indices for rice yield estimation across phenological stages (2022–2023); Table S3. Random Forest parameter tuning (number of trees and variables) for rice yield prediction using vegetation indices at different phenological stages of 2022; Table S4. Random Forest parameter tuning (number of trees and variables) for rice yield prediction at the dough stage (2023); Table S5. Random Forest parameter tuning (number of trees and variables) for rice yield prediction at the milk stage (2022–2023).

Author Contributions: Conceptualization, J.Q.-M. and L.R.-F.; methodology, I.J.-E.; validation, J.Q.-M.; investigation, J.H.-M., D.Q.-T. and E.P.-V.; resources, J.H.-M. and D.Q.-T.; data curation, J.H.-M.; writing—original draft preparation, I.J.-E. and J.H.-M.; writing—review and editing, J.H.-M. and J.Q.-M.; supervision, L.R.-F. and A.T.-R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Program for Scientific Research and Advanced Studies (PROCIENCIA–Peru) under the project "ECOSMART RICE: New precision technological tools with remote sensors for a sustainable rice production system with reduced water consumption, lower greenhouse gas emissions, and higher yield, for the benefit of farmers in the Lambayeque region" (Contract No. PE501086540-2024-PROCIENCIA).

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Acknowledgments: The support of the ECOSMART RICE project (PROCIENCIA–Peru, Contract No. PE501086540-2024-PROCIENCIA) is gratefully acknowledged.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Islam, M.D.; Di, L.; Qamer, F.M.; Shrestha, S.; Guo, L.; Lin, L.; Mayer, T.J.; Phalke, A.R. Rapid Rice Yield Estimation Using Integrated Remote Sensing and Meteorological Data and Machine Learning. *Remote Sens.* **2023**, *15*, 2374. [CrossRef]
2. Mallareddy, M.; Thirumalaikumar, R.; Balasubramanian, P.; Naseeruddin, R.; Nithya, N.; Mariadoss, A.; Eazhilkrishna, N.; Choudhary, A.K.; Deiveegan, M.; Subramanian, E.; et al. Maximizing Water Use Efficiency in Rice Farming: A Comprehensive Review of Innovative Irrigation Management Technologies. *Water* **2023**, *15*, 1802. [CrossRef]

3. Goigochea-Pinchi, D.; Justino-Pinedo, M.; Vega-Herrera, S.S.; Sanchez-Ojanasta, M.; Lobato-Galvez, R.H.; Santillan-Gonzales, M.D.; Ganoza-Roncal, J.J.; Ore-Aquino, Z.L.; Agurto-Piñarreta, A.I. Yield Prediction Models for Rice Varieties Using UAV Multispectral Imagery in the Amazon Lowlands of Peru. *AgriEngineering* **2024**, *6*, 2955–2969. [CrossRef]
4. Quille-Mamani, J.A.; Ruiz, L.A.; Ramos-Fernández, L. Multiseasonal Analysis of Rice Crop Yield Prediction with Sentinel-2 Time Series and UAV Imagery in Lambayeque (Peru). *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2024**, *XLVIII-3-2024*, 421–426. [CrossRef]
5. Fernandez-Beltran, R.; Baidar, T.; Kang, J.; Pla, F. Rice-Yield Prediction with Multi-Temporal Sentinel-2 Data and 3D CNN: A Case Study in Nepal. *Remote Sens.* **2021**, *13*, 1391. [CrossRef]
6. Ibrahim, G.R.F.; Rasul, A.; Abdullah, H. Assessing How Irrigation Practices and Soil Moisture Affect Crop Growth through Monitoring Sentinel-1 and Sentinel-2 Data. *Environ. Monit. Assess.* **2023**, *195*, 1262. [CrossRef] [PubMed]
7. Huber, F.; Yushchenko, A.; Stratmann, B.; Steinhage, V. Extreme Gradient Boosting for Yield Estimation Compared with Deep Learning Approaches. *Comput. Electron. Agric.* **2022**, *202*, 107346. [CrossRef]
8. Impollonia, G.; Croci, M.; Amaducci, S. Upscaling and Downscaling Approaches for Early Season Rice Yield Prediction Using Sentinel-2 and Machine Learning for Precision Nitrogen Fertilisation. *Comput. Electron. Agric.* **2024**, *227*, 109603. [CrossRef]
9. Nazir, A.; Ullah, S.; Saqib, Z.A.; Abbas, A.; Ali, A.; Iqbal, M.S.; Hussain, K.; Shakir, M.; Shah, M.; Butt, M.U. Estimation and Forecasting of Rice Yield Using Phenology-Based Algorithm and Linear Regression Model on Sentinel-II Satellite Data. *Agriculture* **2021**, *11*, 1026. [CrossRef]
10. Chen, J.; Yu, T.; Cherney, J.H.; Zhang, Z. Optimal Integration of Optical and SAR Data for Improving Alfalfa Yield and Quality Traits Prediction: New Insights into Satellite-Based Forage Crop Monitoring. *Remote Sens.* **2024**, *16*, 734. [CrossRef]
11. Andrade, C.; Fonseca, A.; Santos, J.A.; Bois, B.; Jones, G.V. Historic Changes and Future Projections in Köppen–Geiger Climate Classifications in Major Wine Regions Worldwide. *Climate* **2024**, *12*, 94. [CrossRef]
12. Franch, B.; San Bautista, A.; Fita, D.; Rubio, C.; Tarrazó-Serrano, D.; Sánchez, A.; Skakun, S.; Vermote, E.; Becker-Reshef, I.; Uris, A. Within-Field Rice Yield Estimation Based on Sentinel-2 Satellite Data. *Remote Sens.* **2021**, *13*, 4095. [CrossRef]
13. Fan, L.; Fang, S.; Fan, J.; Wang, Y.; Zhan, L.; He, Y. Rice Yield Estimation Using Machine Learning and Feature Selection in Hilly and Mountainous Chongqing, China. *Agriculture* **2024**, *14*, 1615. [CrossRef]
14. Choudhary, K.; Shi, W.; Dong, Y.; Paringer, R. Random Forest for Rice Yield Mapping and Prediction Using Sentinel-2 Data with Google Earth Engine. *Adv. Space Res.* **2022**, *70*, 2443–2457. [CrossRef]
15. Roy, D.P.; Li, J.; Zhang, H.K.; Yan, L. Best Practices for the Reprojection and Resampling of Sentinel-2 Multi Spectral Instrument Level 1C Data. *Remote Sens. Lett.* **2016**, *7*, 1023–1032. [CrossRef]
16. Li, Z.; Zhang, H.K.; Roy, D.P.; Yan, L.; Huang, H. Sharpening the Sentinel-2 10 and 20 m Bands to Planetscope-0 3 m Resolution. *Remote Sens.* **2020**, *12*, 2406. [CrossRef]
17. Yang, K.; Luo, Y.; Li, M.; Zhong, S.; Liu, Q.; Li, X. Reconstruction of Sentinel-2 Image Time Series Using Google Earth Engine. *Remote Sens.* **2022**, *14*, 4395. [CrossRef]
18. Fernandes, R.; Hong, G.; Brown, L.A.; Dash, J.; Harvey, K.; Kalimipalli, S.; MacDougall, C.; Meier, C.; Morris, H.; Shah, H.; et al. Not Just a Pretty Picture: Mapping Leaf Area Index at 10 m Resolution Using Sentinel-2. *Remote Sens. Environ.* **2024**, *311*, 114269. [CrossRef]
19. Raphiphan, Y.; Khetkeeree, S.; Liangrocapart, S. Sharpening the Sentinel-2A Infrared Bands via Image Residual Optimization. In Proceedings of the 19th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, ECTI-CON 2022, Prachuap Khiri Khan, Thailand, 24–27 May 2022. [CrossRef]
20. Panagiotopoulou, A.; Grammatikopoulos, L.; Kalousi, G.; Charou, E. Sentinel-2 and SPOT-7 Images in Machine Learning Frameworks for Super-Resolution. In *Pattern Recognition. ICPR International Workshops and Challenges*; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Cham, Switzerland, 2021; Volume 12667 LNCS, pp. 462–476. [CrossRef]
21. Montero, D.; Aybar, C.; Mahecha, M.D.; Martinuzzi, F.; Söchting, M.; Wieneke, S. A Standardized Catalogue of Spectral Indices to Advance the Use of Remote Sensing in Earth System Research. *Sci. Data* **2023**, *10*, 197. [CrossRef]
22. Tiruneh, G.; Tsegaye Meshesha, D.; Adgo, E.; Tsunekawa, A.; Haregeweyn, N.; Almaw Fenta, A.; Yilak Alemayehu, T.; Mulualem, T.; Fekadu, G.; Demissie, S.; et al. Mapping Crop Yield Spatial Variability Using Sentinel-2 Vegetation Indices in Ethiopia. *Arab. J. Geosci.* **2023**, *16*, 631. [CrossRef]
23. Panek, E.; Gozdowski, D.; Stępień, M.; Samborski, S.; Ruciński, D.; Buszke, B. Within-Field Relationships between Satellite-Derived Vegetation Indices, Grain Yield and Spike Number of Winter Wheat and Triticale. *Agronomy* **2020**, *10*, 1842. [CrossRef]
24. Htun, A.M.; Shamsuzzoha, M.; Ahamed, T. Rice Yield Prediction Model Using Normalized Vegetation and Water Indices from Sentinel-2A Satellite Imagery Datasets. *Asia-Pac. J. Reg. Sci.* **2023**, *7*, 491–519. [CrossRef]
25. Cavalaris, C.; Megoudi, S.; Maxouri, M.; Anatolitis, K.; Sifakis, M.; Levizou, E.; Kyparissis, A. Modeling of Durum Wheat Yield Based on Sentinel-2 Imagery. *Agronomy* **2021**, *11*, 1486. [CrossRef]

26. Al-Gaadi, K.A.; Hassaballa, A.A.; Tola, E.; Kayad, A.G.; Madugundu, R.; Alblewi, B.; Assiri, F. Prediction of Potato Crop Yield Using Precision Agriculture Techniques. *PLoS ONE* **2016**, *11*, e0162219. [CrossRef] [PubMed]
27. Zhang, F.; Hassanzadeh, A.; Letendre, P.; Kikkert, J.; Pethybridge, S.; van Aardt, J. Enhancing Snap Bean Yield Prediction through Synergistic Integration of UAS-Based LiDAR and Multispectral Imagery. *Comput. Electron. Agric.* **2025**, *230*, 109923. [CrossRef]
28. Yewle, A.D.; Mirzayeva, L.; Karakuş, O. Multi-Modal Data Fusion and Deep Ensemble Learning for Accurate Crop Yield Prediction. *Remote Sens. Appl.* **2025**, *38*, 101613. [CrossRef]
29. Wang, J.; Zhang, Y.; Han, F.; Shi, Z.; Zhao, F.; Zhang, F.; Pan, W.; Zhang, Z.; Cui, Q. Estimation of Canopy Chlorophyll Content of Apple Trees Based on UAV Multispectral Remote Sensing Images. *Agriculture* **2025**, *15*, 1308. [CrossRef]
30. Biplov Bhandari, Y.; Mayer, T. Comparing Deep Learning Models for Rice Mapping in Bhutan Using High Resolution Satellite Imagery. *arXiv* **2024**, arXiv:2406.07482. [CrossRef]
31. Niedbała, G.; Hegazi, E.H.; Samak, A.A.; Yang, L.; Huang, R.; Huang, J. Prediction of Soil Moisture Content from Sentinel-2 Images Using Convolutional Neural Network (CNN). *Agronomy* **2023**, *13*, 656. [CrossRef]
32. Benzougagh, B.; Meshram, S.G.; El Fellah, B.; Mastere, M.; Dridri, A.; Sadkaoui, D.; Mimich, K.; Khedher, K.M. Combined Use of Sentinel-2 and Landsat-8 to Monitor Water Surface Area and Evaluated Drought Risk Severity Using Google Earth Engine. *Earth Sci. Inform.* **2022**, *15*, 929–940. [CrossRef]
33. Ramos, J.; Di-Laccio, J.L.; Uzuriaga, J.; Bonomo, N. Remote Sensing of Cyanobacterias with Sentinel-2 in the Salto Grande. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2024**, *XLVIII-2/W6-2024*, 45–50. [CrossRef]
34. Zhang, J.; Pan, Y.; Tao, X.; Wang, B.; Cao, Q.; Tian, Y.; Zhu, Y.; Cao, W.; Liu, X. In-Season Mapping of Rice Yield Potential at Jointing Stage Using Sentinel-2 Images Integrated with High-Precision UAS Data. *Eur. J. Agron.* **2023**, *146*, 126808. [CrossRef]
35. Bazezew, M.N.; Belay, A.T.; Guda, S.T.; Kleinn, C. Developing Maize Yield Predictive Models from Sentinel-2 MSI Derived Vegetation Indices: An Approach to an Early Warning System on Yield Fluctuation and Food Security. *PFG—J. Photogramm. Remote Sens. Geoinf. Sci.* **2021**, *89*, 535–548. [CrossRef]
36. Angulo, L.; Pamboukian, S. Spectral Behavior of Maize, Rice, Soy, and Oat Crops Using Multi-Spectral Images from Sentinel-2. In Proceedings of the 5th Brazilian Technology Symposium, Campinas, Brazil, 22–24 October 2019; Springer: Cham, Switzerland, 2021; Volume 202, pp. 327–336. [CrossRef]
37. Hassanzadeh, A.; Zhang, F.; van Aardt, J.; Murphy, S.P.; Pethybridge, S.J. Broadacre Crop Yield Estimation Using Imaging Spectroscopy from Unmanned Aerial Systems (UAS): A Field-Based Case Study with Snap Bean. *Remote Sens.* **2021**, *13*, 3241. [CrossRef]
38. Kayad, A.; Sozzi, M.; Gatto, S.; Marinello, F.; Pirotti, F. Monitoring Within-Field Variability of Corn Yield Using Sentinel-2 and Machine Learning Techniques. *Remote Sens.* **2019**, *11*, 2873. [CrossRef]
39. Chanev, M.; Kamenova, I.; Dimitrov, P.; Filchev, L. Evaluation of Sentinel-2 Deep Resolution 3.0 Data for Winter Crop Identification and Organic Barley Yield Prediction. *Remote Sens.* **2025**, *17*, 957. [CrossRef]
40. Li, F.; Miao, Y.; Chen, X.; Sun, Z.; Stueve, K.; Yuan, F. In-Season Prediction of Corn Grain Yield through PlanetScope and Sentinel-2 Images. *Agronomy* **2022**, *12*, 3176. [CrossRef]
41. Wan, L.; Ryu, Y.; Dechant, B.; Lee, J.; Zhong, Z.; Feng, H. Improving Retrieval of Leaf Chlorophyll Content from Sentinel-2 and Landsat-7/8 Imagery by Correcting for Canopy Structural Effects. *Remote Sens. Environ.* **2024**, *304*, 114048. [CrossRef]
42. Skakun, S.; Vermote, E.; Franch, B.; Roger, J.C.; Kussul, N.; Ju, J.; Masek, J. Winter Wheat Yield Assessment from Landsat 8 and Sentinel-2 Data: Incorporating Surface Reflectance, Through Phenological Fitting, into Regression Yield Models. *Remote Sens.* **2019**, *11*, 1768. [CrossRef]
43. Panumonwatee, G.; Choosumrong, S.; Pampasit, S.; Premprasit, R.; Nemoto, T.; Raghavan, V. Machine Learning Technique for Carbon Sequestration Estimation of Mango Orchards Area Using Sentinel-2 Data. *Carbon Res.* **2025**, *4*, 33. [CrossRef]
44. Escolà, A.; Badia, N.; Arnó, J.; Martínez-Casasnovas, J.A. Using Sentinel-2 Images to Implement Precision Agriculture Techniques in Large Arable Fields: First Results of a Case Study. *Adv. Anim. Biosci.* **2017**, *8*, 377–382. [CrossRef]
45. Radočaj, D.; Jurišić, M. A Phenology-Based Evaluation of the Optimal Proxy for Cropland Suitability Based on Crop Yield Correlations from Sentinel-2 Image Time-Series. *Agriculture* **2025**, *15*, 859. [CrossRef]
46. Parida, P.K.; Somasundaram, E.; Krishnan, R.; Radhamani, S.; Sivakumar, U.; Parameswari, E.; Raja, R.; Shri Rangasami, S.R.; Sangeetha, S.P.; Gangai Selvi, R. Unmanned Aerial Vehicle-Measured Multispectral Vegetation Indices for Predicting LAI, SPAD Chlorophyll, and Yield of Maize. *Agriculture* **2024**, *14*, 1110. [CrossRef]
47. Soccolini, A.; Vizzari, M. Predictive Modelling of Maize Yield Using Sentinel 2 NDVI. In *Computational Science and Its Applications—ICCSA 2023 Workshops*; Springer: Cham, Switzerland, 2023; Volume 14107 LNCS, pp. 327–338. [CrossRef]
48. Pinna, D.; Sozzi, M.; Marinello, F. Assessing Silage Maize Yield and Quality Variability Using NIRS, Sentinel-2, and PlanetScope Multispectral Imagery: A Precision Agriculture Approach. *Remote Sens. Agric.* **2024**, *13191*, 182–189. [CrossRef]
49. Chaiyana, A.; Hanchoowong, R.; Srihanu, N.; Prasanchum, H.; Kangrang, A.; Hormwichian, R.; Kaewplang, S.; Koedsin, W.; Huete, A. Leveraging Remotely Sensed and Climatic Data for Improved Crop Yield Prediction in the Chi Basin, Thailand. *Sustainability* **2024**, *16*, 2260. [CrossRef]

50. Jabed, M.A.; Azmi Murad, M.A. Crop Yield Prediction in Agriculture: A Comprehensive Review of Machine Learning and Deep Learning Approaches, with Insights for Future Research and Sustainability. *Heliyon* **2024**, *10*, e40836. [CrossRef] [PubMed]
51. Shafiee, S.; Lied, L.M.; Burud, I.; Dieseth, J.A.; Alsheikh, M.; Lillemo, M. Sequential Forward Selection and Support Vector Regression in Comparison to LASSO Regression for Spring Wheat Yield Prediction Based on UAV Imagery. *Comput. Electron. Agric.* **2021**, *183*, 106036. [CrossRef]
52. Fawagreh, K.; Gaber, M.M.; Elyan, E. Random Forests: From Early Developments to Recent Advancements. *Syst. Sci. Control Eng.* **2014**, *2*, 602–609. [CrossRef]
53. Ji, F.; Li, F.; Hao, D.; Shiklomanov, A.N.; Yang, X.; Townsend, P.A.; Dashti, H.; Nakaji, T.; Kovach, K.R.; Liu, H.; et al. Unveiling the Transferability of PLSR Models for Leaf Trait Estimation: Lessons from a Comprehensive Analysis with a Novel Global Dataset. *New Phytol.* **2024**, *243*, 111–131. [CrossRef]
54. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
55. Biau, G.; Scornet, E. A Random Forest Guided Tour. *Test* **2015**, *25*, 197–227. [CrossRef]
56. Dhillon, M.S.; Dahms, T.; Kuebert-Flock, C.; Rummler, T.; Arnault, J.; Steffan-Dewenter, I.; Ullmann, T. Integrating Random Forest and Crop Modeling Improves the Crop Yield Prediction of Winter Wheat and Oil Seed Rape. *Front. Remote Sens.* **2022**, *3*, 1010978. [CrossRef]
57. Jhajharia, K.; Sharma, N.V.; Mathur, P. A Machine Learning Model for Crop Yield Prediction Using Remote Sensing Data. *Int. Res. J. Multidiscip. Scope* **2025**, *6*, 577–590. [CrossRef]
58. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. *arXiv* **2016**, arXiv:1603.02754. [CrossRef]
59. Sah, S.; Haldar, D.; Singh, R.N.; Das, B.; Nain, A.S. Rice Yield Prediction through Integration of Biophysical Parameters with SAR and Optical Remote Sensing Data Using Machine Learning Models. *Sci. Rep.* **2024**, *14*, 21674. [CrossRef] [PubMed]
60. Tiwari, V.; Thorp, K.; Tulbure, M.G.; Gray, J.; Kamruzzaman, M.; Krupnik, T.J.; Sankarasubramanian, A.; Ardon, M. Advancing Food Security: Rice Yield Estimation Framework Using Time-Series Satellite Data & Machine Learning. *PLoS ONE* **2024**, *19*, e0309982. [CrossRef]
61. Najafi, P.; Eftekhari, A.; Sharifi, A. Evaluation of Time-Series Sentinel-2 Images for Early Estimation of Rice Yields in South-West of Iran. *Aircr. Eng. Aerosp. Technol.* **2023**, *95*, 741–748. [CrossRef]
62. Soriano-González, J.; Angelats, E.; Martínez-Eixarch, M.; Alcaraz, C. Monitoring Rice Crop and Yield Estimation with Sentinel-2 Data. *Field Crops Res.* **2022**, *281*, 108507. [CrossRef]
63. Sun, C.; Zhang, H.; Xu, L.; Ge, J.; Jiang, J.; Song, M.; Wang, C. Rice Yield Prediction Using Radar Vegetation Indices from Sentinel-1 Data and Multiscale One-Dimensional Convolutional Long- and Short-Term Memory Network Model. *J. Appl. Remote Sens.* **2024**, *18*, 024505. [CrossRef]
64. Pokhariyal, S.; Patel, N.R.; Nain, A.S.; Akarsh, S.G.; Rana, R.S.; Singh, R.K.; Ranjan, R. Evaluating Rice Crop Phenology and Crop Yield in Hilly Region Using Satellite Imagery and Google Earth Engine. *J. Agrometeorol.* **2024**, *26*, 395–400. [CrossRef]
65. Quille-Mamani, J.; Ramos-Fernández, L.; Huanuqueño-Murillo, J.; Quispe-Tito, D.; Cruz-Villacorta, L.; Pino-Vargas, E.; Flores del Pino, L.; Heros-Aguilar, E.; Ángel Ruiz, L. Rice Yield Prediction Using Spectral and Textural Indices Derived from UAV Imagery and Machine Learning Models in Lambayeque, Peru. *Remote Sens.* **2025**, *17*, 632. [CrossRef]
66. Kalecinski, N.I.; Skakun, S.; Torbick, N.; Huang, X.; Franch, B.; Roger, J.C.; Vermote, E. Crop Yield Estimation at Different Growing Stages Using a Synergy of SAR and Optical Remote Sensing Data. *Sci. Remote Sens.* **2024**, *10*, 100153. [CrossRef]
67. Guzman-Lopez, R.R.; Huamanchumo, L.; Fernandez, K.; Cutipa-Luque, O.; Tiahuallpa, Y.; Rojas, H. Sparsity, Regularization and Causality in Agricultural Yield: The Case of Paddy Rice in Peru. *arXiv* **2024**, arXiv:2409.17298. [CrossRef]
68. Alebele, Y.; Wang, W.; Yu, W.; Zhang, X.; Yao, X.; Tian, Y.; Zhu, Y.; Cao, W.; Cheng, T. Estimation of Crop Yield from Combined Optical and SAR Imagery Using Gaussian Kernel Regression. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 10520–10534. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.