remote sensing

MDPI

Article

Rice Yield Prediction Using Spectral and Textural Indices Derived from UAV Imagery and Machine Learning Models in Lambayeque, Peru

Javier Quille-Mamani [1], Lia Ramos-Fernández [2,*], José Huanuqueño-Murillo [2], David Quispe-Tito [2], Lena Cruz-Villacorta [3], Edwin Pino-Vargas [4], Lisveth Flores del Pino [5], Elizabeth Heros-Aguilar [6] and Luis Ángel Ruiz [1]

[1] Geo-Environmental Cartography and Remote Sensing Group (CGAT), Universitat Politècnica de València, Camí de Vera s/n, 46022 Valencia, Spain; jaquille@upv.es (J.Q.-M.); laruiz@cgf.upv.es (L.Á.R.)
[2] Department 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] Department of Territorial Planning and Doctoral Program in Engineering and Environmental Sciences, Universidad Nacional Agraria La Molina, Lima 15024, Peru; lenacruz@lamolina.edu.pe
[4] Department of Civil Engineering, Jorge Basadre Grohmann National University, Tacna 23000, Peru; epinov@unjbg.edu.pe
[5] Center for Research in Chemistry, Toxicology and Environmental Biotechnology, National Agrarian University La Molina, Lima 15024, Peru; lisveth@lamolina.edu.pe
[6] Department of Phytotechnics, National Agrarian University La Molina, Lima 15024, Peru; lizheros@lamolina.edu.pe
* Correspondence: liarf@lamolina.edu.pe

Abstract: Predicting rice yield accurately is crucial for enhancing farming practices and securing food supplies. This research aims to estimate rice yield in Peru's Lambayeque region by utilizing spectral and textural indices derived from unmanned aerial vehicle (UAV) imagery, which offers a cost-effective alternative to traditional approaches. UAV data collection in commercial areas involved seven flights in 2022 and ten in 2023, focusing on key growth stages such as flowering, milk, and dough, each showing significant predictive capability. Vegetation indices like NDVI, SP, DVI, NDRE, GNDVI, and EVI2, along with textural features from the gray-level co-occurrence matrix (GLCM) such as ENE, ENT, COR, IDM, CON, SA, and VAR, were combined to form a comprehensive dataset for model training. Among the machine learning models tested, including Multiple Linear Regression (MLR), Support Vector Machines (SVR), and Random Forest (RF), MLR demonstrated high reliability for annual data with an R^2 of 0.69 during the flowering and milk stages, and an R^2 of 0.78 for the dough stage in 2022. The RF model excelled in the combined analysis of 2022–2023 data, achieving an R^2 of 0.58 for the dough stage, all confirmed through cross-validation. Integrating spectral and textural data from UAV imagery enhances early yield prediction, aiding precision agriculture and informed decision-making in rice management. These results emphasize the need to incorporate climate variables to refine predictions under diverse environmental conditions, offering a scalable solution to improve agricultural management and market planning.

Keywords: vegetation indices (VIs); textural indices (TIs); multiple linear regression (MLR); support vector regression (SVR); random forest (RF); cross-validation; machine learning

Academic Editor: Annamaria Castrignanò

Received: 6 December 2024
Revised: 31 January 2025
Accepted: 11 February 2025
Published: 12 February 2025

Citation: 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. https://doi.org/10.3390/rs17040632

1. Introduction

Rice is the most widely consumed staple food globally, feeding more than 50% of the world's population [1]. It plays a critical role in the diets of Asia, Latin America, and the

Caribbean, while its consumption is significantly increasing in Africa, reflecting a growing trend in global demand [2]. For the 2023/24 period, world rice production is projected to reach 526.2 million tons, representing a 0.4% increase compared to the revised estimate for 2022/23 [3].

In Peru, paddy rice production under continuous flooded irrigation (CF) reached 1.5 million tons in 2023, with the Lambayeque and La Libertad regions being the primary contributors. Yields of 9.51 t ha^{-1} were recorded across over 220,000 hectares nationwide [4]. However, this production entails significant resource consumption, with water usage ranging from 12,000 to 20,000 m^3 ha^{-1} and nitrogen applications reaching 200 to 320 kg N ha^{-1} [5]. These practices pose critical challenges, including water scarcity [6] and inefficiencies in nitrogen management [7], which exacerbate production costs, greenhouse gas emissions, and pest vulnerability, while threatening food security [8]. Addressing these issues requires the development of pre-harvest rice yield prediction models to inform national food policy and resource management strategies [6].

Numerous studies indicate that the pre-panicle emergence stage is optimal for predicting grain yield using vegetation indices (VIs) [9–11]. VIs derived from spectral data have shown strong correlations with crop phenotyping and final yields [12–15]. Specifically, the normalized difference vegetation index (NDVI) and Visible Atmospherically Resistant Index (VARI) demonstrate high accuracy in estimating rice yields at the booting stage [11,16].

High-resolution imagery captured by unmanned aerial vehicles (UAVs) equipped with RGB and multispectral cameras has proven effective for rice yield prediction, utilizing indices such as NDVI, Normalized Difference Water Index (NDWI), Radar Vegetation Index (RVI), and difference vegetation index (DVI) [9,11,17,18]. These methods have been successfully applied in diverse regions, including China [10,11,19], Spain [12], and Malaysia [20]. Furthermore, incorporating canopy structural parameters, such as plant height and canopy cover, alongside VIs enhances predictive accuracy, as demonstrated by Su et al. [15].

Beyond rice, these approaches have shown utility in other crops, including wheat [21], maize [14], and beans [13]. Various VIs, such as the Chlorophyll Index Red Edge (CIrededge), Chlorophyll Index Green (CIgreen), enhanced vegetation index (EVI), two-band enhanced vegetation index (EVI2), and MERIS Terrestrial Chlorophyll Index (MTCI), have also been effectively utilized for yield predictions in rice fields [9,11,15,17,22,23]. However, their sensitivity typically diminishes during the reproductive growth stages [15].

UAV-derived RGB imagery provides detailed textural information, enabling the precise quantification of vegetation parameters [24]. Among the various approaches, textural analysis focusing on the Normalized Difference Texture Index (NDTI) has shown significant potential for biomass estimation in rice crops, outperforming conventional vegetation indices in some cases [11]. In northeastern Colorado, USA, textural characteristics combined with field measurements were effectively used to predict aboveground biomass that could not be captured by spectral indices alone [25]. Moreover, textural data have proven valuable in estimating panicle formation, a critical parameter for field-level crop yield prediction [21].

A widely adopted method for extracting textural features is the gray-level co-occurrence matrix (GLCM), which applies statistical techniques to analyze spatial relationships within an image [24]. The GLCM framework allows for the derivation of 28 distinct textural measures, of which eight are most commonly used: Energy (ENE), Entropy (ENT), Correlation (COR), Homogeneity (HOM), Contrast (CON), Dissimilarity (DIS), Sum Average (SA), and Variance (VAR) [25–29]. These measures are instrumental in enhancing

the understanding of textural patterns, further supporting biomass estimation and yield prediction.

Studies have demonstrated the effectiveness of integrating data from the visible and infrared spectrum with textural information to enhance the accuracy of paddy rice yield predictions [30]. The combination of textural and spectral features in high-resolution imagery has proven valuable for applications such as crop classification modeling, leaf area index estimation, and crop growth monitoring [19,29,31]. Composite indices derived from this integration provide a robust theoretical framework for subsequent crop yield assessments [19,29]. While spectral features primarily describe pixel intensity in each band, textural features capture the spatial distribution of these pixel intensities, offering complementary insights into crop conditions [24].

In this setting, employing hyperspectral and LiDAR sensors on UAVs has been demonstrated to greatly enhance yield prediction across different crops. For instance, in maize, the synergy between hyperspectral images and LiDAR information resulted in R^2 values ranging from 0.75 to 0.79, surpassing the effectiveness of using each sensor independently [32]. Similarly, in bean cultivation, merging these two technologies led to an adjusted R^2 of 0.83 in the harvest period [33]. Furthermore, studies on rice crops emphasized that deploying UAVs equipped with hyperspectral sensor technology markedly boosted the capacity for yield prediction, showcasing an improvement in the precision of the acquired data [34].

In recent years, multispectral and optical remote sensing technologies using UAVs have gained prominence as efficient tools for crop assessment due to their ability to capture high-resolution images in real time [11,15,35]. Despite their advantages, such as superior spatial resolution, UAVs face limitations in cost and their ability to cover extensive areas. Consequently, open-access satellite imagery, including Sentinel and Landsat, with a resolution of 30 m, has emerged as a viable alternative. These satellite platforms have demonstrated reliable accuracy in crop yield predictions, albeit slightly lower than that of UAV imagery. However, satellite imagery is constrained by lower temporal resolution, restricting its capability for continuous, plot-level monitoring [12].

Given these advancements, this study aims to develop a yield prediction model for commercial rice areas in the Lambayeque region. By combining spectral indices and textural measurements obtained from UAVs, the research seeks to optimize agricultural practices in arid ecosystems. The proposed approach highlights the potential of integrating remote sensing technologies for real-time decision-making, ultimately contributing to improved agricultural management and food security.

2. Materials and Methods

2.1. Study Sites

The study was conducted in the Lambayeque region, focusing on commercial areas well known for rice farming. This area is situated in the province of Ferreñafe and includes five distinct zones, where 37 plots were analyzed in 2022 and 35 plots in 2023 (refer to Figure 1), as outlined in Table 1. The climate in the research area is categorized as BWh, which signifies a warm desert climate under the Köppen classification system [36]. The region experiences an average yearly temperature of 22 °C and receives an annual rainfall of 20 mm.

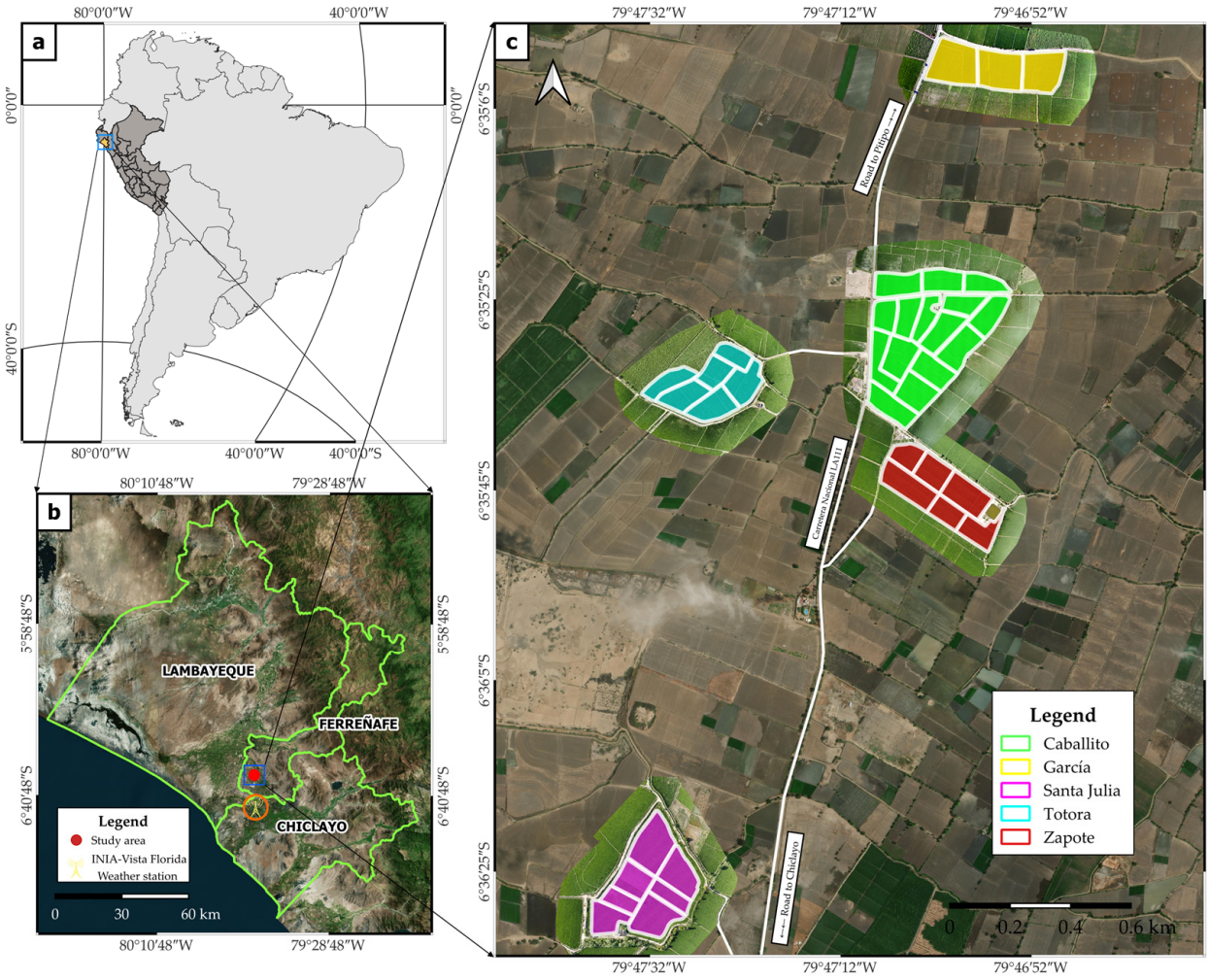

Figure 1. Study area: (**a**) geographical location of Peru; (**b**) Lambayeque region; and (**c**) commercial zones: Caballito, García, Santa Julia, Totora, and Zapote.

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
Garcia	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

2.2. Meteorological Characteritics

The meteorological data for the study area during the months of rice crop development, from January to June 2022 and 2023, include relative humidity (%), maximum temperature (°C), minimum temperature (°C), precipitation (mm), and wind speed (m s^{-1}). These data were recorded at the automatic weather station of INIA-Vista Florida (Figure 2).

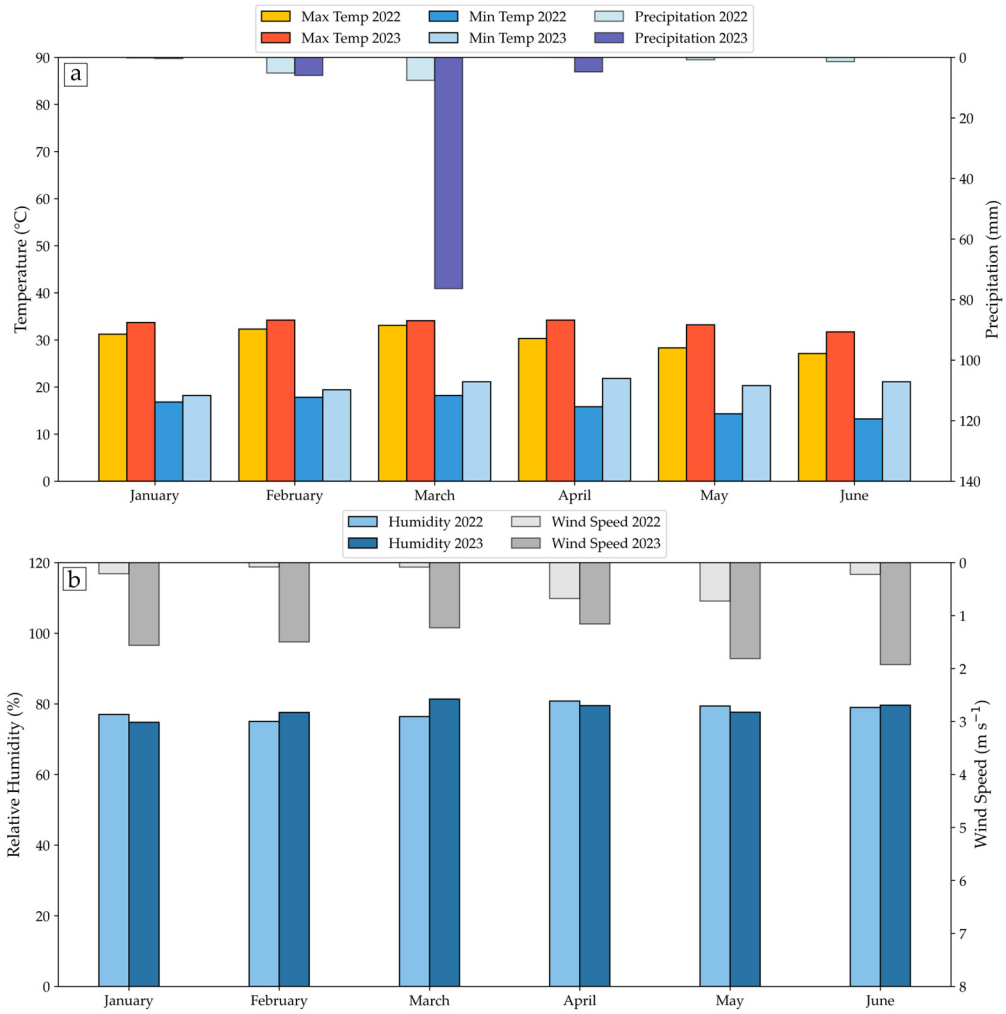

Figure 2. Meteorological variables recorded during the rice growing season in 2022 and 2023: (**a**) maximum temperature (°C), minimum temperature (°C), and precipitation (mm); (**b**) relative humidity (%) and wind speed (m s^{-1}). These data were collected at the automatic weather station of INIA-Vista Florida.

2.3. Soil Characterization

Table 2 provides an overview of the soil's physical properties in the commercial rice fields analyzed in the study. Sampling was carried out at depth of 25 cm in five distinct locations: Caballito, García, Santa Julia, Totora, and Zapote. The data include information on soil texture, bulk density, true density, field capacity, and permanent wilting point for each location. The findings reveal variations in soil characteristics, such as bulk density, ranging from 1.40 g cm^{-3} to 1.53 g cm^{-3}, and field capacity values varying between 17.79% and 29.50%. These variations highlight the diverse soil conditions present in the commercial rice cultivation areas.

Table 2. Soil physical characteristics of commercial rice fields.

Zones	Soil Texture	Bulk Density (g cm^{-3})	Real Density (g cm^{-3})	Field Capacity (%)	Permanent Wilting Point (%)
Caballito	Sandy loam	1.45	2.33	24.24 ± 0.96	22.54 ± 1.33
García	Sandy loam	1.47	2.47	23.31 ± 0.73	12.33 ± 0.27
Santa Julia	Clay loam	1.46	2.56	20.38 ± 0.53	11.79 ± 0.37
Totora	Clay loam	1.53	2.6	17.79 ± 1.01	10.24 ± 0.50
Zapote	Clay loam	1.4	2.53	29.50 ± 1.96	16.58 ± 0.53

2.4. Crop Management

During the 2022 and 2023 growing seasons, work was conducted in five commercial zones: Caballito, Totora, Garcia, Santa Julia and Zapote. In 2022, 37 subplots were evaluated, while in 2023, the number was reduced to 35 due to salinity problems. The subplots ranged from 5 to 12 hectares and the rice varieties Tinajones, Capoteña, Puntilla, Mallares, and Pakamuros (Table 1), known for their tolerance to water stress, were grown. Tinajones, with a semi-early cycle, has recorded yields of up to 14.0–15.0 t ha^{-1}, while Capoteña, with an intermediate cycle, has reached more than 12.0–13.0 t ha^{-1} under commercial conditions. Puntilla, Mallares, and Pakamuros have achieved up to 13.0 t ha^{-1} under optimal agronomic management. Transplanting was carried out 35 days after sowing (DDS) with a planting frame of 25 × 25 cm and 2 seedlings per stroke, under a continuous flooding (CI) system. In Zapote, fertilization was 265 N-72 P-50 K, while in the other zones, 263 N-92 P-75 K was used. Fertilizer application was distributed at three points: before transplanting (5% N, 85% P and 100% K) and at 18 DDS (42% N and 15% P), and the rest of the nitrogen was applied in equal parts during the tillering and cotton point stages.

Crop phenology was assessed using growing degree days (GDD), a highly effective index for tracking and understanding crop development [37]. This method provides a quantitative measure of heat accumulated over time, which is crucial for identifying key milestones in crop growth. Figure 3 depicts the phenological cycle of the Capoteña rice crop and Table 3 shows the flights during the phenological cycle of the study plots.

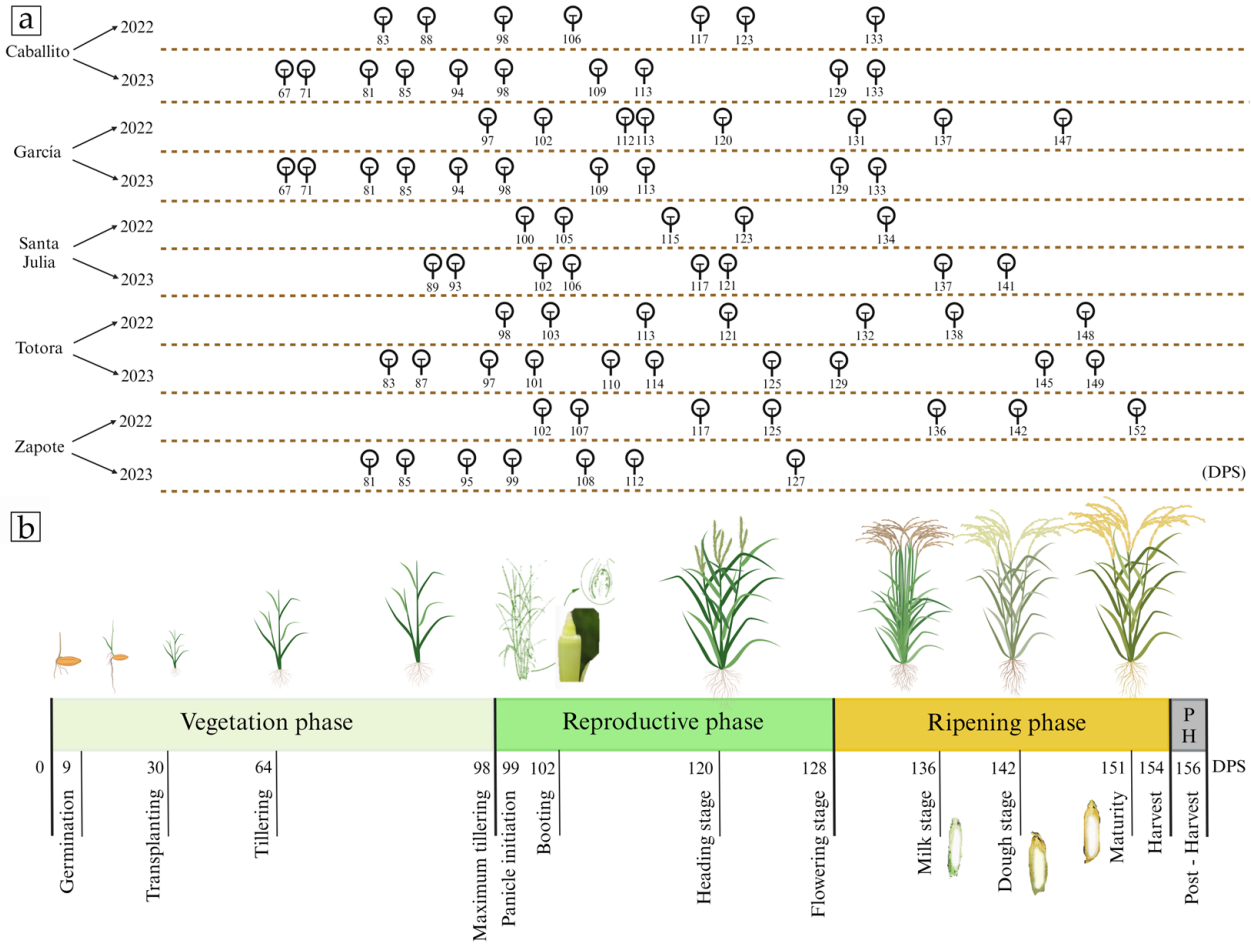

Figure 3. (**a**) Flights carried out in the commercial areas; (**b**) phenology of the Capoteña variety according to days post sowing (DPS).

Table 3. UAV flights for the corresponding two-year growth phases in the commercial areas.

Variety	Flight Date and Days Post Sowing (DPS)									
	Tinajones y Capoteña (Caballito)		Tinajones (Garcia)		Mallares (Santa Julia)		Puntilla (Totora)		Pakamuros (Zapote)	
Growth Stage	**2022**	**2023**	**2022**	**2023**	**2022**	**2023**	**2022**	**2023**	**2022**	**2023**
Tillering	*	06-March (67)	*	06-March (67)	*	*	*	*	*	*
Panicle Initiation	*	10-March (71)	*	10-March (71)	*	*	*	*	*	*
Heading Stage	*	20-March (81)	*	20-March (81)	*	*	*	*	*	06-March (81)
Heading Stage	17-March (83)	24-March (85)	*	24-March (85)	*	20-March (89)	*	06-March (83)	*	10-March (85)
Flowering Stage	22-March (88)	02-April (94)	17-March (97)	02-April (94)	*	24-March (93)	*	20-March (97)	*	20-March (95)
Flowering Stage	01-April (98)	06-April (98)	22-March (102)	06-April (98)	17-March (100)	02-April (102)	17-March (98)	24-March (101)	17-March (102)	24-March (99)
Milk Stage	09-April (106)	17-April (109)	01-April (112)	17-April (109)	22-March (105)	06-April (106)	22-Mar (103)	02-April (110)	22-March (107)	02-April (108)
Milk Stage	20-April (117)	21-April (113)	09-April (120)	21-April (113)	01-April (115)	17-April (117)	01-April (113)	06-April (114)	01-April (117)	06-April (112)
Dough Stage	26-April (123)	07-May (129)	20-April (131)	07-May (129)	09-April (123)	21-April (121)	09-April (121)	17-April (125)	09-April (125)	21-April (127)
Dough Stage	06-May (133)	11-May (133)	26-April (137)	11-May (133)	20-April (134)	07-May (137)	20-April (132)	21-April (129)	20-April (136)	31-May (167)
Maturity	*	*	06-May (147)	*	*	11-May (141)	26-April (138)	07-May (145)	26-April (142)	*
Harvest	*	*	*	*	*	*	06-May (148)	11-May (149)	06-May (152)	*

* No drone flights were conducted during the 2022 and 2023 growing seasons for the corresponding phenological stages.

2.5. Data Acquisition and Processing

Figure 4 presents the workflow designed for predicting rice yield based on data collected during the 2022 and 2023 growing seasons. The process begins with the acquisition of multispectral and RGB images using UAVs (unmanned aerial vehicles), complemented by ground measurements. As indicated in Table 3, images were captured at a minimum frequency of one flight every one to two weeks, between 10:00 a.m. and 2:00 p.m., to ensure optimal lighting conditions for data acquisition. From these images, various spectral indices are computed, including the normalized difference vegetation index (NDVI), normalized difference red-edge vegetation index (NDRE), normalized difference vegetation index green (GNAVI), simple index (SP), difference vegetation index (DVI), and enhanced vegetation index 2 (EVI2). Additionally, texture indices such as energy (ENE), entropy (ENT), correlation (COR), inverse difference moment (IDM), contrast (CON), mean sum (SA), and variance (VAR) are extracted to accurately capture the conditions and variability of the crops.

The extracted indices are then processed using Python, where feature selection methods, such as sequential forward selection, are applied to refine the model inputs. Machine learning techniques, including Multiple Linear Regression (MLR), Support Vector Machine Regression (SVR) with linear and radial basis function kernels, and Random Forest (RF), are employed to develop predictive models. A "grid search" procedure is used to fine-tune hyperparameters for optimal model performance.

The validation of the model was conducted using Leave-One-Out Cross-Validation (LOOCV) instead of traditional validation due to the limited sample size, allowing for optimal utilization of available data and generating more robust performance estimates. Furthermore, LOOCV minimizes information loss and reduces variability compared to other methods, ensuring a more reliable and accurate assessment. In the final stage of the workflow, the selected features and the most effective model are integrated to predict rice yield. This approach leverages advanced remote sensing data and machine learning techniques to produce accurate and reliable yield forecasts.

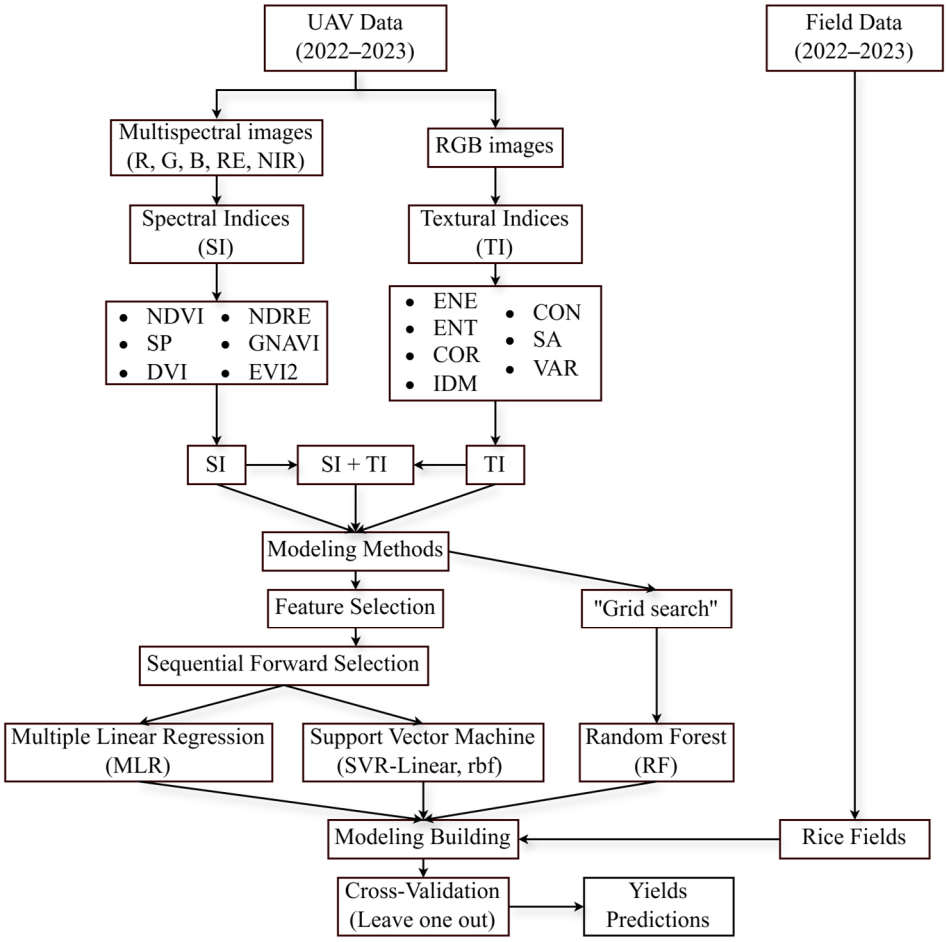

Figure 4. Flow diagram of the methodology followed in this study.

2.5.1. Multispectral and RGB Imaging with an Unmanned Aerial Vehicle (UAV)

The aerial photographs were obtained using a DJI Matrice 300-RTK UAV from Shenzhen, China. This drone weighs 8.37 kg and measures 810 × 670 × 430 mm (length × width × height). It features a Global Navigation Satellite System (GNSS) receiver along with real-time kinematic (RTK) positioning systems that work with the Global Positioning System (GPS). The UAV boasts a hovering accuracy of ±0.1 m vertically and horizontally in D-RTK mode, when there is no wind. A distinct multispectral sensor was used for each survey year, as shown in Figure 5. In 2022, the sensor employed was the Micasense RedEdge-MX, produced by MicaSense, Inc. in Seattle, WA, USA, which includes a downward light sensor (DLS2). In 2023, the Parrot Sequoia, developed by Parrot S.A. in France, was used. Both sensors are equipped with a sunlight sensor that assesses the light intensity for their specific spectral bands, as detailed in Table 4. The captured images underwent radiometric, atmospheric, and geometric adjustments. For radiometric calibration, a calibrated reflectance panel specific to each sensor and a sunlight sensor that automatically modifies the readings based on ambient light were utilized [5]. The flight plan was crafted considering the climatic conditions and the characteristics of the study area, as outlined in Table 5.

Figure 5. Flight platform and sensors. (**a**) DJI Matric 300 RTK, (**b**) Micasense RedEdge-MX multispectral sensor, and (**c**) Parrot Sequoia multispectral sensor, together with their respective calibration panels.

Table 4. Wavelength bands of multispectral sensors.

Band Name	MicaSense RedEdge-Mx	Parrot Sequoia
Blue	459–491 nm	–
Green	546.5–573.5 nm	480–520 nm
Red	661–675 nm	640–680 nm
Red-edge	711–723 nm	730–740 nm
Near-infrared	813.5–870.5 nm	770–810 nm

Table 5. Flight plans for commercial areas.

Zones	Resolution (cm Pixel^{-1})	Frontal and Lateral Overlap (%)	Velocity (m s^{-1})	Height (m)	Time of Flight	Area (m^2)
Caballito	7	85 × 80	8.6	120	12′51″	181,000
García	7	85 × 80	8.6	120	5′	76,000
Santa Julia	7	85 × 80	8.6	120	7′ 50″	109,000
Totora	7	85 × 80	8.6	120	5′41″	74,000
Zapote	7	85 × 80	8.6	120	19′54″	72,000

The images that were captured were processed using the Pix4Dmapper Pro software (Pix4D S.A., Prilly, Switzerland), specifically version 4.4.12. This process was executed through a series of sequential steps. Initially, the flight parameters were inputted, and the camera specifications were configured, allowing the software to identify and process the images accurately. Following this, the georeferencing of ground control points (GCPs) placed on the ground was conducted, ensuring that at least four control points were present for every monitored area. The next step involved adjusting the mesh's density and texture to create a point cloud that represented the surface. Afterward, the elevation model and orthomosaic were developed. Finally, a georeferencing error analysis was conducted, ensuring that the margin of error did not exceed 5 cm. To achieve this level of accuracy, RTK or an optimal distribution of a sufficient number of GCPs was utilized. If the margin

of error exceeded this threshold, the process was restarted from the beginning, potentially increasing the number of images used in the georeferencing process [5].

2.5.2. Measurement of Rice Grain Yield

To assess the yield of rice, the initial step involved determining the moisture content of the dehusked grains using a WILE-55 moisture meter manufactured by Farmcomp Oy in Finland. This instrument provides moisture readings within the range of 8% to 30%, maintaining an accuracy of ±0.5%, and functions effectively between temperatures of 0 °C and 40 °C. Measurements were taken once the grains achieved a moisture level of around 16%, ensuring that the moisture content was standardized to approximately 14% during the weighing process. The commercial fields were harvested using mechanical methods, occurring between days 140 and 166 of the crop growth cycle (DPS) in 2022, and between days 139 and 157 DPS in 2023, contingent upon the specific subplot being assessed (refer to Figure 6).

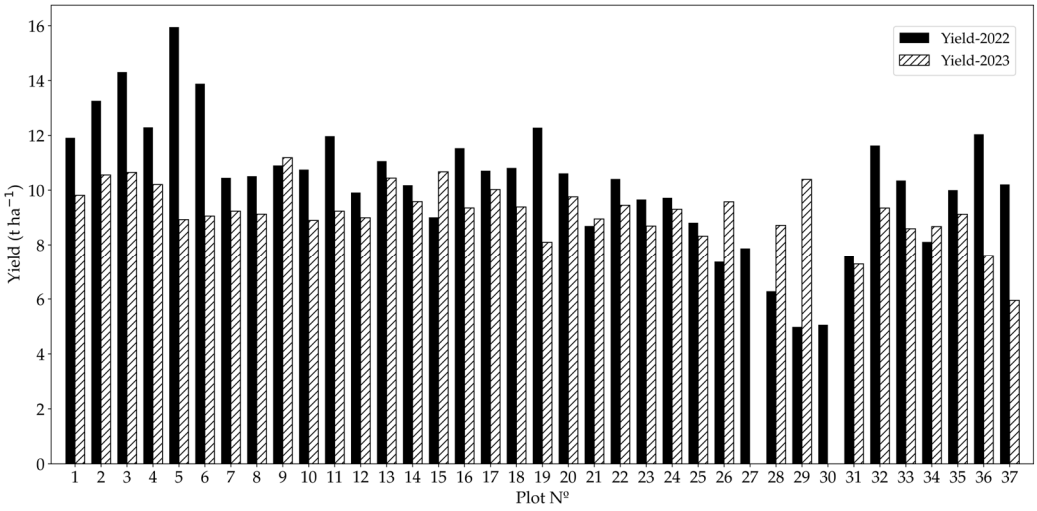

Figure 6. Rice yield data in tons per hectare (t ha^{-1}) in commercial fields of Ferreñafe for the years 2022 and 2023.

2.5.3. Spectral and Textural Index Calculations

VIs were extracted from multispectral orthomosaics obtained via UAV technology, utilizing the green, NIR, red, and red-edge bands. Six specific VIs were chosen for the study—NDVI, SP, DVI, NDRE, GNDVI, and EVI2 (Table 6)—all of which are frequently used in research to forecast crop yield. To derive TIs, RGB orthomosaics were analyzed using the gray-level co-occurrence matrix (GLCM) approach [38,39]. This research utilized seven well-established TIs: ENE, ENT, COR, HOM, CON, SA, and VAR (Table 7). The extraction of texture based on GLCM was conducted with QGIS version 3.16.3-Hannover, using the texture-GRASS plugin (https://grass.osgeo.org/grass-stable/manuals/r.texture.html, accessed on 10 October 2024) [40]. Both VIs and TIs served as input variables for the model predicting yield.

Table 6. Applied spectral indices for yield evaluation in rice.

Spectral Indices	Calculation Formula	Sources
Normalized Difference Vegetation Index (NDVI)	$NDVI = \frac{NIR-Red}{NIR+Red}$	[16,18,19,41]
Simple Ratio (SP)	$SP = \frac{NIR}{Red}$	[1,11,19,29]
Difference Vegetation Index (DVI)	$DVI = NIR - Red$	[11,17,19]
Normalized Difference Red Edge (NDRE)	$NDRE = \frac{NIR-Red\ Edge}{NIR+Red\ Edge}$	[13–15]
Green Normalized Difference Vegetation Index (GNDVI)	$GNDVI = \frac{NIR-Green}{NIR+Green}$	[14,15]
Enhanced Vegetation Index 2 (EVI 2)	$EVI2 = 2.4 \times \frac{NIR-Red}{NIR+Red+1}$	[15,17,21]

Table 7. Textural indices applied for rice yield evaluation.

Texture Indices	Calculation Formula	Sources
Angular Second Moment (ENE)	$\sum_i \sum_j p(i,j)^2$	[1,29]
Entropy (ENT)	$-\sum_i \sum_j p(i,j)\log(p((i,j)))$	[25]
Correlation (COR)	$\frac{\sum_i \sum_j (ij(p(i,j)-\mu_x\mu_y}{\sigma_x\sigma_y}$	[25,31]
Inverse Difference Moment (IDM)	$\sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$	[31]
Contrast (CON)	$\sum_{n=0}^{N_8-1} n^2 \left\{ s\sum_{i=1}^{N_8-1} \sum_{j=1}^{N_8-1} p(i,j) \right\}$	[1,29]
Sum Average (SA)	$\sum_{i=2}^{2N} i \cdot p_{x+y}(i)$	[1]
Variance (VAR)	$\sum_i \sum_j (i-u)^2 p(i,j)$	[25,31]

$p(i,j)$ represents the probability of each pixel pair (i,j), where i and j are the gray levels within the window and coordinates in the co-occurrence matrix. Ng denotes the number of unique gray levels in the quantized image, corresponding to the original image's gray value range. μ and σ represent the mean and standard deviation of $p(i,j)$, respectively.

2.6. Modeling Methods

2.6.1. Multiple Linear Regression with Sequential Forward Selection (SFS-MLR)

MLR is commonly used to create calibration models; however, it may not be suitable for datasets with numerous independent variables [42]. To enhance its performance, linear models that can be interpreted through multiple linear regression with sequential forward selection (SFS-MLR) analysis were utilized. This modeling approach was executed using the Mlxtend Sequential Feature Selector module (https://github.com/rasbt/mlxtend, accessed 10 March 2024) [43] within the Python programming environment (Version 3.13).

2.6.2. Support Vector Machine-Based Sequential Forward Selection Regression (SFS-SVR)

The sequential forward selection (SFS) algorithm begins by employing a single feature to model the data using the given model [44]. A significant benefit of SFS is its capacity to automatically choose a subset of more pertinent features, enhancing computational efficiency and lowering generalization error by eliminating irrelevant or noisy features. This is especially advantageous for algorithms lacking support for regularization [45]. Paired with the SVR [44], the variable selection process initially utilizes SFS. The procedure initiates with an empty set and progressively incorporates the variable that most enhances model performance [46]. This continues until the optimal input combination, based on an evaluation function, is discovered. In this research, SFS-SVR with 'linear' and 'rbf' kernels, set with C parameters of 1, a tolerance of 0.001, and epsilon of 0.1, served as the induction model. Mean square error was used as the criterion for minimization, and resampling techniques were applied at each iteration to stabilize feature rankings. The

variables chosen by SFS were then used as inputs for regression modeling. The processes of variable selection and regression modeling were executed using the Python packages 'Mlxtend' and 'Sklearn'.

2.6.3. Random Forest Regression (RFR)

RF is a supervised learning algorithm that combines multiple decision trees (ntree) generated from random samples of predictor variables [46]. Furthermore, RF is robust against collinearity among variables, which enhances prediction accuracy and minimizes overfitting [47]. In this research, three key parameters were fine-tuned using a "grid search" approach—the number of rounds (ranging from 1 to 150), the maximum depth of the trees (between 3 and 20), and the maximum number of branching features (from 2 to 16)—to forecast rice crop yield utilizing VIs and TIs. The model was developed using the Python 'Sklearn' library.

2.7. Model Validation

To evaluate the precision of the models in forecasting rice yield values at the plot level, the coefficient of determination (R^2), root-mean-square error (RMSE), and relative-mean-square error (rRMSE) were employed. These measures enable the estimation of errors in rice yield predictions. To ensure a comprehensive evaluation in this research, LOOCV was utilized [44]. The calculation of these metrics was performed 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 jth measured and predicted values of rice yield, respectively; $\bar{Y}$ is the average of measured yield.

3. Results

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

Figure 7 presents data on the number of plots assessed and the R^2 obtained by simple linear regression between yield and VIs and TIs throughout various rice growth stages in 2022, 2023, and a combined analysis of both years. The prediction models were developed by considering the average values of the flowering, milk, and dough stages, as these are crucial for ensuring accurate predictions.

In Figure 7a, there is a noticeable increase in the number of plots evaluated as the phenological stages progress, with the milk and dough stages in 2023 showing the highest numbers. This increase in plot numbers contributed to more reliable yield predictions, especially during these significant stages. Figure 7b illustrates the distribution of the R^2 for each growth stage in 2022 and 2023. In the 2022 season, flowering stage 2 proved to be the most reliable for model development, with consistently higher R^2 values compared to other stages. Conversely, in 2023, flowering stage 1 was identified as the best, showing a similar trend of high R^2 values. When data from both years were combined, the flowering

stage continued to demonstrate strong predictive reliability, emphasizing its importance in yield prediction models.

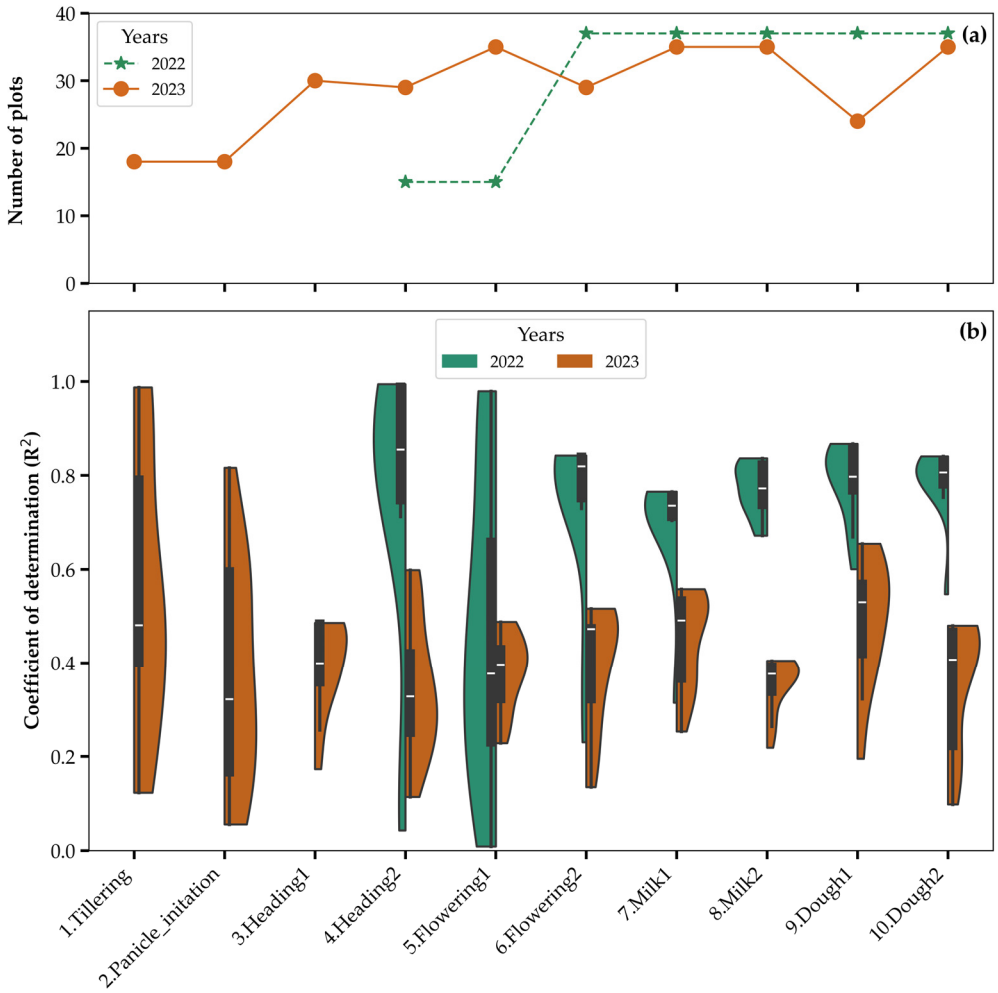

Figure 7. Coefficient of determination (R^2) of vegetation indices (VIs) and textural indices (TIs) in relation to measured rice yield during phenological stages. (**a**) Number of plots evaluated for each phenological stage in 2022 and 2023. (**b**) Distribution of R^2 values across phenological stages for 2022 and 2023.

It is important to note the absence of data for the tillering, panicle initiation, and heading stages in 2023. This lack of information is due to the limited availability of images during these phases, either due to unfavorable weather conditions, operational constraints in data acquisition, or insufficient image quality. Despite this limitation, the results suggest that post-tillering stages, particularly flowering and milk, remain crucial for yield prediction.

These findings emphasize the importance of choosing the correct phenological stages for model calibration and validation, along with the impact of sample size (number of plots) on the consistency and dependability of predictions. The flowering stages consistently proved effective across both years, highlighting their significance in yield modeling.

3.2. Correlation Analysis Between Multiple-Period Vegetation Indices (VIs) and Textural Indices (TIs) and Rice Yield

These results underscore the potential of vegetation and texture indices as predictive tools for assessing rice yield, with variations observed based on phenological stage and the period of analysis. Table 8 summarizes the correlation coefficients between VIs, TIs,

and rice yield across the 2022 and 2023 growing seasons, as well as the combined dataset (2022–2023).

Table 8. Specific values and significance levels of correlation coefficients for spectral and textural characteristics with rice yield in 2022, 2023 and both years combined (2022–2023).

Spectral and Textural Characteristics	Correlation Coefficients (r)								
	Flowering			Milk			Dough		
	2022	2023	2022–2023	2022	2023	2022–2023	2022	2023	2022–2023
NDVI	0.01 ns	0.15 ns	−0.14 ns	0.55 ***	0.27 ns	0.40 ***	0.59 ***	0.02 ns	0.42 ***
NDRE	0.06 ns	0.46 **	0.11 ns	0.63 ***	0.35 *	0.57 ***	0.71 ***	0.14 ns	0.56 ***
GNDVI	−0.16 ns	0.28 ns	−0.24 *	0.61 ***	0.44 **	0.45 ***	0.73 ***	0.12 ns	0.49 ***
SP	0.02 ns	0.11 ns	−0.13 ns	0.47 **	0.26 ns	0.38 ***	0.49 **	0.03 ns	0.43 ***
DVI	0.50 **	0.01 ns	−0.04 ns	0.69 ***	−0.17 ns	−0.04 ns	0.60 ***	0.01 ns	0.18 ns
EVI2	0.41 *	0.01 ns	−0.04 ns	0.68 ***	−0.13 ns	0.03 ns	0.59 ***	−0.01 ns	0.28 *
ASM	0.11 ns	−0.09 ns	0.22 ns	0.43 **	−0.40 *	0.40 ***	0.59 ***	−0.15 ns	0.52 ***
CONTR	−0.49 **	−0.29 ns	−0.29 *	−0.07 ns	0.40 *	−0.09 ns	−0.34 *	0.20 ns	−0.08 ns
CORR	−0.25 ns	−0.01 ns	−0.32 *	−0.17 ns	−0.31 ns	−0.32 **	−0.59 ***	−0.09 ns	−0.47 ***
ENTR	−0.12 ns	0.07 ns	−0.23 *	−0.32 ns	0.34 *	−0.34 **	−0.64 ***	0.19 ns	−0.49 ***
IDM	0.15 ns	−0.04 ns	0.26 *	0.24 ns	−0.50 **	0.30 *	0.42 *	−0.24 ns	0.36 **
SA	−0.16 ns	0.23 ns	−0.24 *	−0.21 ns	0.04 ns	−0.33 **	−0.45 **	−0.14 ns	−0.48 ***
VAR	−0.50 **	−0.30 ns	−0.29 *	−0.07 ns	0.39 *	−0.10 ns	−0.36 *	0.20 ns	−0.09 ns

Note: ns is not significant at the $p \geq 0.05$ level; * is significantly correlated at the $p < 0.05$ level; ** is significantly correlated at the $p < 0.01$ level; and *** is significantly correlated at the $p < 0.001$ level.

In 2022, VIs proved to be more effective yield predictors compared to TIs, particularly during the milk and dough stages. In the flowering stage, most indices showed no significant correlations, except for DVI (r = 0.50, $p < 0.01$) and EVI2 (r = 0.41, $p < 0.05$), which exhibited a moderate relationship with yield. During the milk stage, NDRE (r = 0.63, $p < 0.001$), GNDVI (r = 0.61, $p < 0.001$), DVI (r = 0.69, $p < 0.001$), and EVI2 (r = 0.68, $p < 0.001$) were the most reliable for yield estimation. In the dough stage, NDRE (r = 0.71, $p < 0.001$) and GNDVI (r = 0.73, $p < 0.001$) maintained their strong predictive capacity. On the other hand, textural indices displayed greater variability in their correlations, with significant negative relationships in the flowering and dough stages, such as CONTR (r = −0.49, $p < 0.01$) and VAR (r = −0.50, $p < 0.01$), suggesting that higher texture heterogeneity might be associated with lower yields. Overall, the milk and dough stages were the most relevant for yield estimation, with NDRE and GNDVI emerging as the most consistent spectral indices, while textural indices had a lesser influence and, in some cases, a negative correlation with crop productivity.

In contrast, in the 2023 season, the relationship between VIs and yield was weaker compared to 2022, indicating that environmental or management factors may have influenced the predictive ability of spectral reflectance on rice productivity. In the flowering stage, most indices showed no significant correlations, except for NDRE (r = 0.46, $p < 0.01$) which exhibited a moderate association with yield. In the milk stage, NDRE (r = 0.35, $p < 0.05$) and GNDVI (r = 0.44, $p < 0.01$) were the most strongly related to yield, although their values were lower than those in 2022. In the dough stage, VIs showed no significant correlations, suggesting that in this year, reflectance at this stage was not closely linked to rice yield. Regarding TIs, negative correlations were observed in the milk stage, particularly in ASM (r = −0.40, $p < 0.05$) and IDM (r = −0.50, $p < 0.01$), suggesting that lower texture variability was associated with higher yields. In the dough stage, the relationships between textural indices and productivity were inconsistent. Overall, the milk stage was the phenological stage with the highest predictive capacity for yield, with NDRE and GNDVI as the most reliable spectral indices, although with lower correlations than in 2022. Textural indices had less influence on yield estimation and, in some cases, showed negative correlations, suggesting that greater texture homogeneity could be associated with higher crop productivity.

For the combined dataset (2022–2023), VIs showed a stronger relationship with yield compared to TIs, with the milk and dough stages being the most suitable phases for yield prediction. In the flowering stage, most spectral indices were not significant, and some, such as NDVI (r = −0.14, ns) and GNDVI (r = −0.24, $p < 0.05$), even showed negative values, confirming that this stage is not optimal for yield estimation. In the milk stage, NDRE (r = 0.57, $p < 0.001$), GNDVI (r = 0.45, $p < 0.001$), and NDVI (r = 0.40, $p < 0.001$) were the most reliable predictors, although with a slight decrease in correlation magnitude compared to 2022, suggesting an effect of interannual variability. In the dough stage, NDRE (r = 0.56, $p < 0.001$) and GNDVI (r = 0.49, $p < 0.001$) remained the strongest predictors, although with slightly lower values than in 2022. On the other hand, TIs exhibited more variable relationships with yield, with significant negative correlations in the milk stage, such as ENTR (r = −0.34, $p < 0.01$) and CORR (r = −0.32, $p < 0.01$), indicating that lower texture variability could be associated with higher yields. Overall, the combined analysis of 2022 and 2023 confirms that the milk and dough stages were the most suitable phenological stages for yield prediction, with NDRE and GNDVI as the most consistent spectral indices, although with a slight reduction in correlation strength due to interannual variability.

These findings highlight the dynamic relationships between spectral and textural indices and rice yield, with substantial interannual variability. The results underscore the value of integrating multi-season analyses and combining spectral and textural metrics to enhance the robustness of predictive models under diverse environmental conditions.

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

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

The prediction process employed MLR, support vector regression with a linear kernel (SVR-linear), and support vector regression with a radial basis function kernel (SVR-rbf), utilizing the SFS technique. Three distinct data scenarios were analyzed: VIs, TIs, and a combination of both (VIs + TIs). Figure 8 displays the variables selected that correlated with the optimal model performance, emphasizing those that achieved the lowest RMSE. Meanwhile, Figures S1–S9 present the selected variables and their performance for the individual years 2022 and 2023, as well as the combined period of 2022–2023. Furthermore, Tables S1–S3 provide a detailed summary of the performance metrics for each model across all scenarios and time periods assessed.

Figure 9 presents the results of the best-performing models for predicting grain yield (t ha^{-1}) across different phenological stages of rice cultivation. In 2022, the analysis was conducted with a sample of $n = 37$ for the flowering, milk, and dough stages, while in 2023, due to a reduction in the number of subplots, the study used $n = 35$ for the same stages. For the combined period 2022–2023, a consolidated sample of $n = 72$ was analyzed.

These models, crafted using MLR and SVR-rbf, employed the SFS technique to determine the best combination of VIs and TIs. Each subplot of the figure showcases the models' predictive capabilities through R^2_CV values and root-mean-square error from cross-validation (RMSE_CV), highlighting their effectiveness in accurately estimating grain yield. Figure 9 underscores the selected features that were instrumental in reducing prediction errors and enhancing model accuracy across different years and phenological phases.

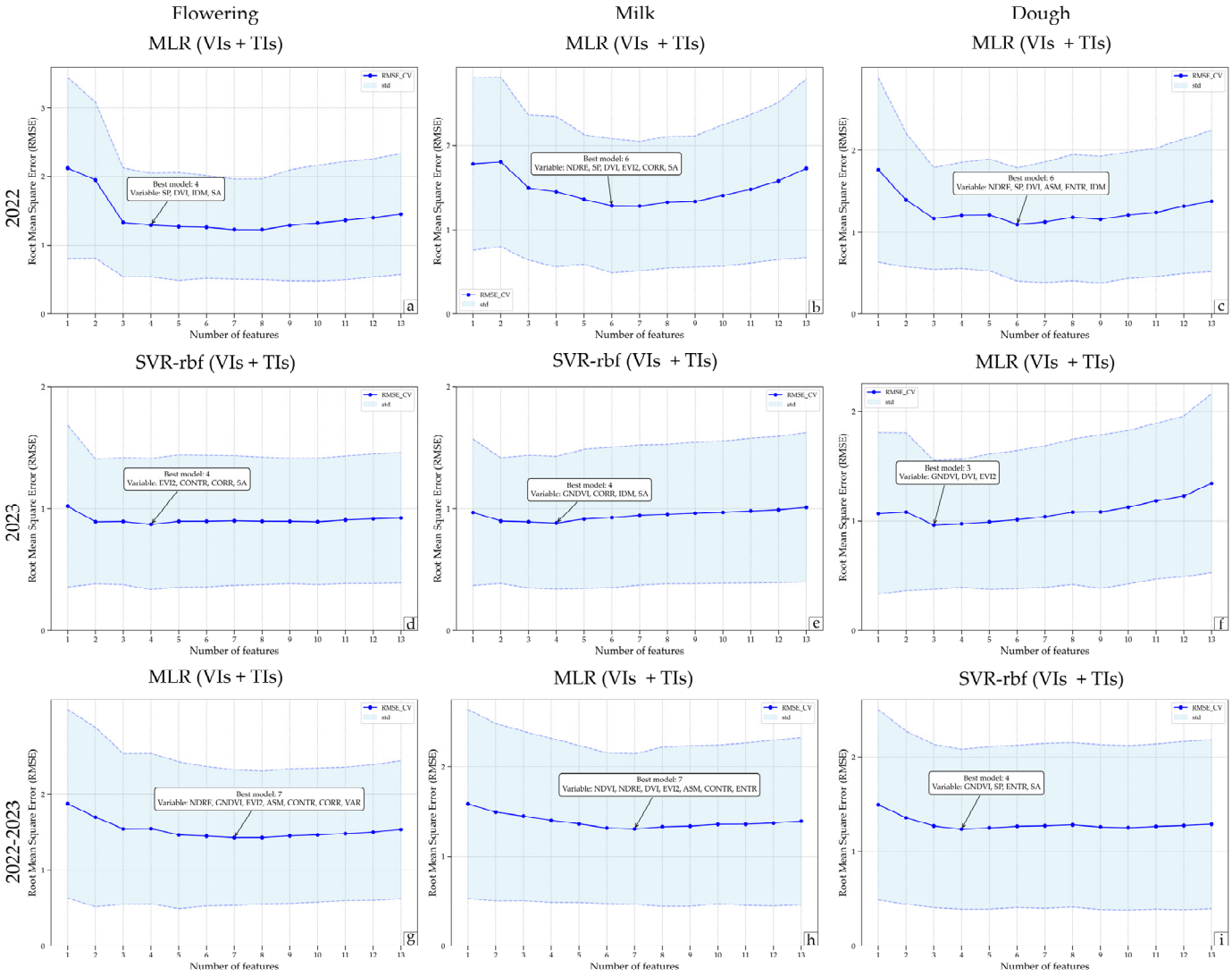

Figure 8. The optimal results from Sequential Feature Selection for Multiple Linear Regression (MLR) and Support Vector Regression (SVR) models using vegetation indices (VIs), texture indices (TIs), and their combination (VIs + TIs) across the flowering (**a**,**d**,**g**), milk (**b**,**e**,**h**), and dough (**c**,**f**,**i**) stages for the years 2022 (**a**–**c**), 2023 (**d**–**f**), and the combined period of 2022–2023 (**g**–**i**).

3.3.2. Random Forest (RF)-Based Yield Modeling

An RF model was created to estimate rice yield during three key phenological stages: flowering, milk and dough. In 2022, the analysis was conducted with a sample of $n = 37$ per phenological stage, while in 2023, due to the reduction in the number of subplots, $n = 35$ was used. For the combined period 2022–2023, a consolidated sample of $n = 72$ was used. The model optimization was conducted using a grid search approach, modifying both the number of trees and the predictor variables related to VIs, TIs, and their combination (VIs + TIs). The goal was to minimize the RMSE and enhance prediction accuracy.

In the flowering phase, the RF model achieved its best performance with 41 trees using VIs, 146 trees with TI, and 121 trees when both indices were combined (VIs + TIs), as shown in Figure 10. The selection process pinpointed three variables for VIs, 1 for TIs, and four for the combined indices. These results highlight the necessity of using diverse indices to capture the complexity of variables during this vital stage of crop development.

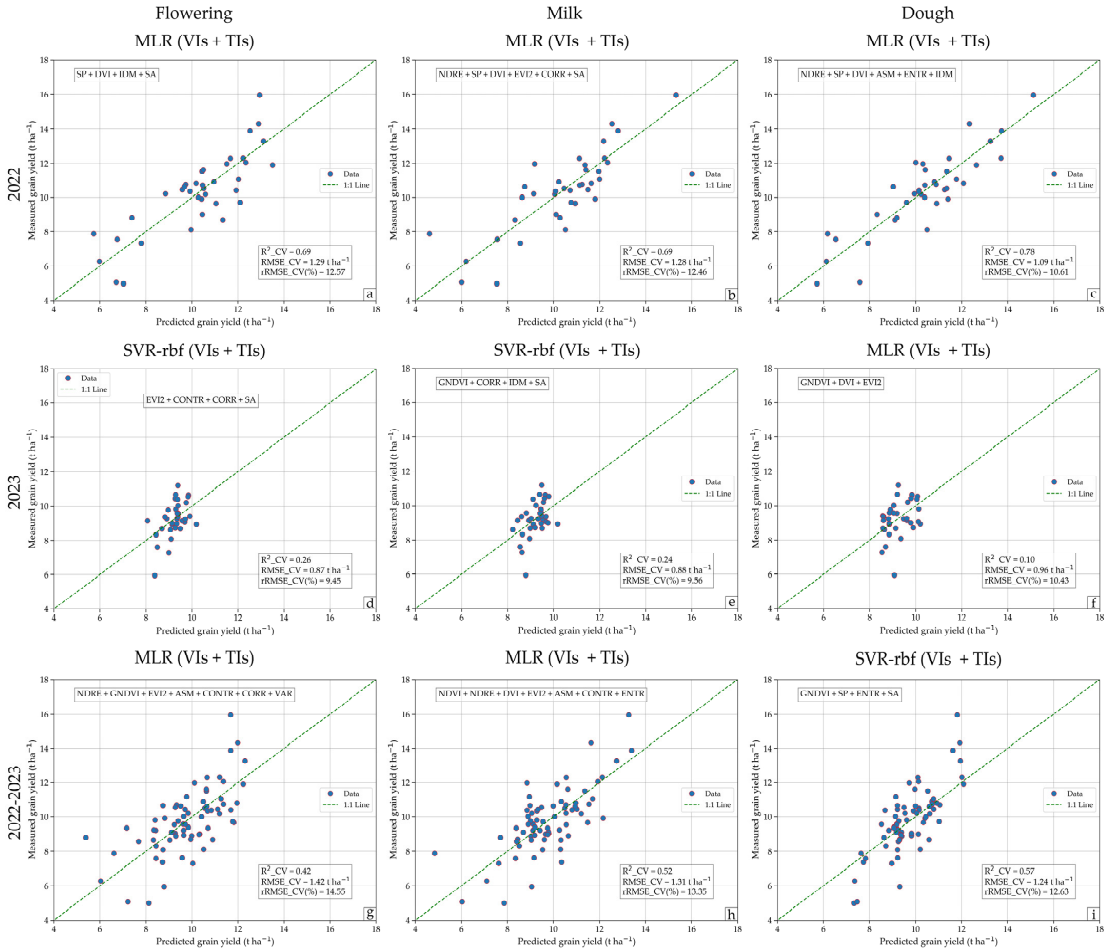

Figure 9. Predicted versus measured grain yield for Multiple Linear Regression (MLR) and Support Vector Regression (SVR) models using vegetation indices (VIs), texture indices (TIs), and their combination (VIs + TIs) across the flowering (**a**,**d**,**g**), milk (**b**,**e**,**h**), and dough (**c**,**f**,**i**) stages for the years 2022 (**a**–**c**), 2023 (**d**–**f**), and the combined period 2022–2023 (**g**–**i**).

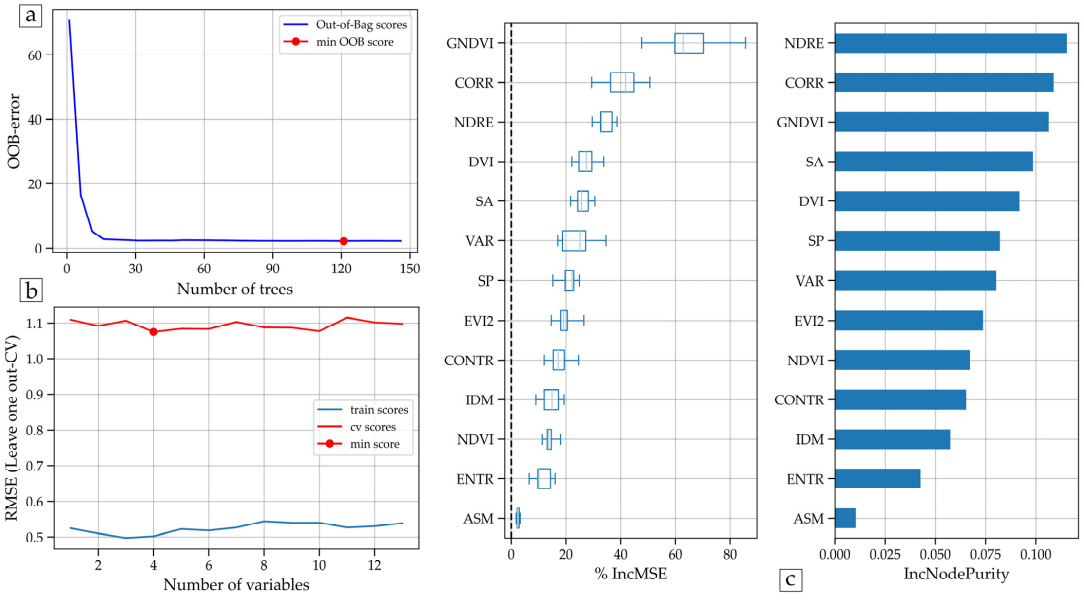

Figure 10. Random Forest (RF) model for rice yield estimation during the flowering stage (2022–2023) using vegetation (VIs) and textural indices (TIs): (**a**) out-of-bag error (OOB), (**b**) variable selection via LOOCV (RMSE), and (**c**) predictor importance.

During the milk phase, the RF model needed 101 trees for VIs, 116 for TIs, and 46 for the combined indices, as detailed in Figure 11. The chosen variables were three for VIs, four for TIs, and one for the combination. At this stage, TIs had a greater predictive influence, emphasizing their importance during this growth period.

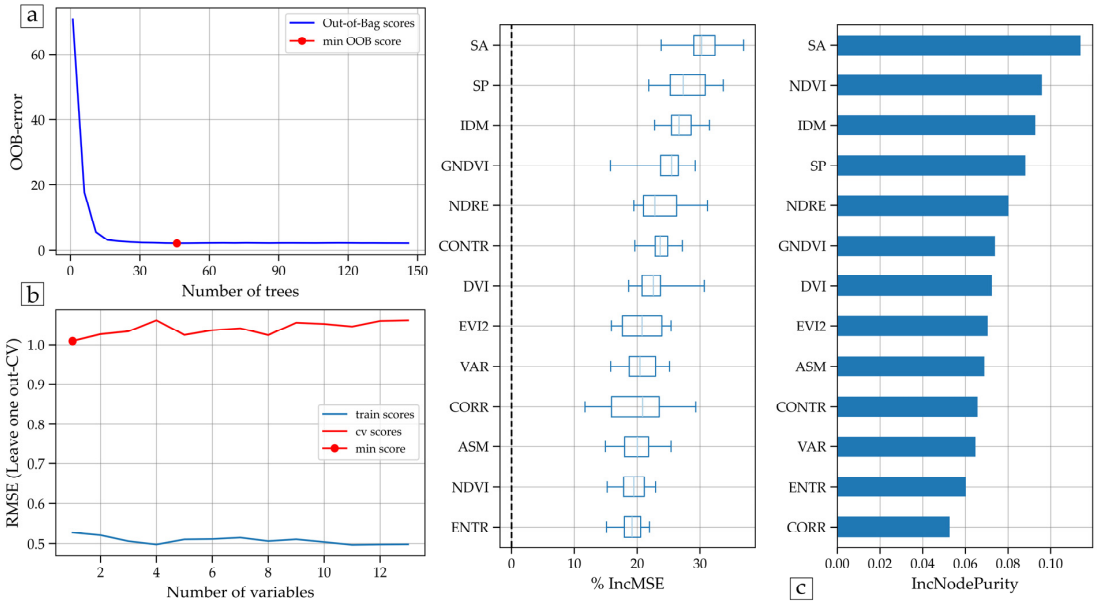

Figure 11. Random Forest (RF) model for rice yield estimation during the milk stage (2022–2023) using vegetation (VIs) and textural indices (TIs): (**a**) out-of-bag error (OOB), (**b**) variable selection via LOOCV (RMSE), and (**c**) predictor importance.

For the dough phase, the model performed better with 146 trees using VIs, 41 trees with TIs, and 71 trees with the combination of both indices, as illustrated in Figure 12. Here, two variables were selected for VIs, seven for TIs, and three for the combined indices. These findings suggest that TIs are particularly crucial during the maturation phase, significantly enhancing model accuracy.

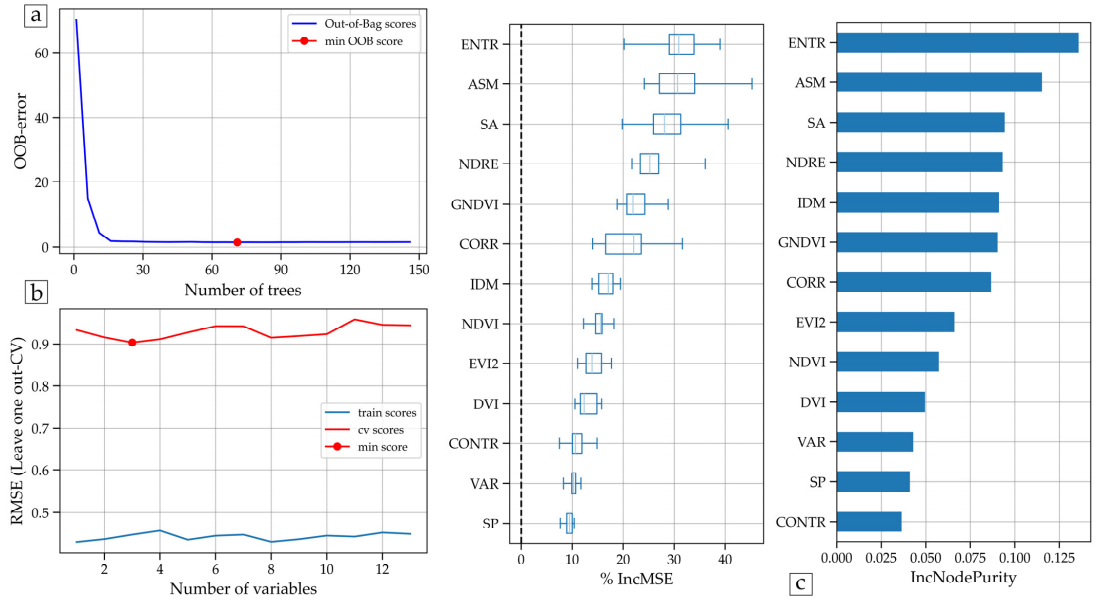

Figure 12. Random Forest (RF) model for rice yield estimation during the dough stage (2022–2023) using vegetation (VIs) and textural indices (TIs): (**a**) out-of-bag error (OOB), (**b**) variable selection via LOOCV (RMSE), and (**c**) predictor importance.

The specific configurations and selected variables are detailed in Tables S4–S6, offering further insights into analysis and replication of the results obtained.

Figure 13 illustrates the predicted grain yield outcomes versus the measured yield, incorporating various combinations of explanatory variables and crop phenological stages. RF models are presented, utilizing VIs, TIs, and a combination of both (VIs + TIs), for the phenological stages of flowering, milk, and dough for the 2022 and 2023 seasons, as well as for the combined cycles of 2022–2023.

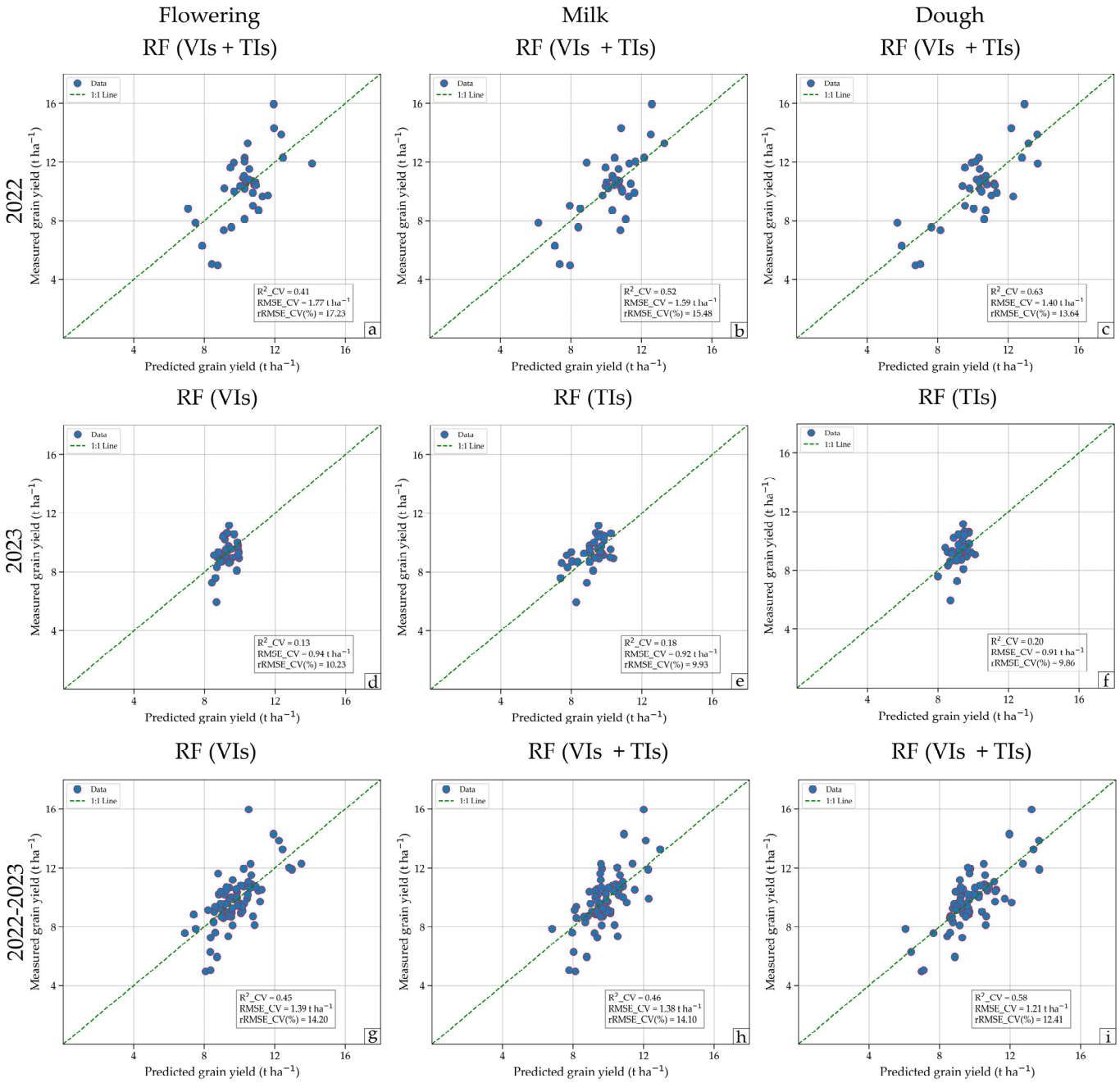

Figure 13. Predicted versus measured grain yield for Random Forest (RF) models using vegetation indices (VIs), texture indices (TIs), and their combination (VIs + TIs) across the flowering (**a**,**d**,**g**), milk (**b**,**e**,**h**), and dough (**c**,**f**,**i**) stages for the years 2022 (**a**–**c**), 2023 (**d**–**f**), and the combined period 2022–2023 (**g**–**i**).

In the flowering stage of 2022–2023 (Figure 13g), the RF model based on VIs achieved the best outcome, with a coefficient of determination (R2) of 0.45, an RMSE_CV of 1.39 t

ha^{-1}, and an RMSE-CV% of 14.20%. During the milk stage of 2022 (Figure 13b), a notable result was obtained with an R^2_CV of 0.52 and an error similar to that of the flowering stage; additionally, for the same year, in the dough stage (Figure 13c), the performance increased, with an R^2_CV of 0.63 and an RMSE_CV of 1.40 t ha^{-1}.

In summary, the RF models based on the combination of VIs and TIs demonstrated the best predictive performance across all the phenological stages analyzed, especially during flowering in 2022 and for the combined 2022–2023 cycle in the flowering stage, as shown in Figure 13.

3.3.3. Performance of the Yield Prediction Models

Table 9 shows the effectiveness of the yield prediction models by employing three different machine learning techniques: MLR, SVR, and RF. These models were validated using the LOOCV method. The outcomes of the regression analysis are depicted in Figures 9 and 13, while the criteria for selecting variables are outlined in Sections 2.6.1–2.6.3. In the RF model, the out-of-bag (OOB) method was used to determine the most impactful variables, with the findings detailed in Tables S4–S6.

Table 9. Assessment of machine learning model performance in forecasting grain yield at various growth stages for 2022, using Leave-One-Out Cross-Validation (CV).

Machine Learning Models	Validation (Leave-One-Out CV)								
	Flowering			Milk			Dough		
	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 (0.80)**	**12.94 (7.75)**	**0.58**	**1.50 (0.86)**	**14.54 (8.36)**	**0.75**	**1.16 (0.62)**	**11.29 (6.02)**
Support Vector Machine (SVR-linear)	0.52	1.61 (1.06)	15.59 (10.31)	0.44	1.72 (1.03)	16.71 (10.05)	0.66	1.35 (0.86)	13.09 (8.38)
Support Vector Machine (SVR-rbf)	0.47	1.68 (1.15)	16.35 (11.14)	0.39	1.81 (1.19)	17.55 (11.60)	0.52	1.59 (1.02)	15.46 (9.91)
Random Forest (RF)	0.32	1.90 (1.36)	18.45 (13.21)	0.43	1.75 (1.01)	16.95 (9.83)	0.50	1.64 (1.02)	15.91 (9.94)
Texture Index (TI)									
Multiple Linear Regression (MLR)	0.15	2.13 (1.27)	20.68 (12.31)	0.22	2.04 (1.25)	19.77 (12.15)	0.59	1.48 (0.92)	14.34 (8.96)
Support Vector Machine (SVR-linear)	0.20	2.07 (1.36)	20.07 (13.21)	0.25	2.00 (1.31)	19.45 (12.77)	**0.60**	**1.45 (0.98)**	**14.12 (9.53)**
Support Vector Machine (SVR-rbf)	**0.40**	**1.79 (1.19)**	**17.34 (11.59)**	**0.32**	**1.91 (1.21)**	**18.51 (11.75)**	0.51	1.61 (1.08)	15.68 (10.52)
Random Forest (RF)	0.26	1.98 (1.19)	19.22 (11.53)	0.15	2.13 (1.43)	20.72 (13.92)	0.50	1.64 (1.04)	15.91 (10.13)
all (IV+TI)									
Multiple Linear Regression (MLR)	**0.69**	**1.29 (0.76)**	**12.57 (7.33)**	**0.69**	**1.28 (0.80)**	**12.46 (7.75)**	**0.78**	**1.09 (0.69)**	**10.61 (6.68)**
Support Vector Machine (SVR-linear)	0.50	1.63 (0.99)	15.86 (9.61)	0.63	1.40 (1.00)	13.61 (9.75)	0.72	1.21 (0.76)	11.77 (7.39)
Support Vector Machine (SVR-rbf)	0.51	1.62 (1.10)	15.76 (10.64)	0.48	1.67 (1.08)	16.17 (10.50)	0.62	1.41 (0.98)	13.74 (9.49)
Random Forest (RF)	0.41	1.77 (1.05)	17.23 (10.22)	0.52	1.59 (1.07)	15.48 (10.40)	0.63	1.40 (0.85)	13.64 (8.30)

Note: Values in bold represent the best performance achieved by the machine learning models at each phenological stage (Flowering, Milk, and Dough).

Based on the information presented in Table 9, the MLR model, which incorporates both VIs and TIs, demonstrated the most exceptional performance in forecasting grain yield across different growth stages in 2022. It achieved the highest R^2_CV values at every stage. Specifically, the model attained an R^2_CV of 0.69 during the flowering stage, an R^2_CV of 0.69 during the milk stage, and an R^2_CV of 0.78 during the dough stage. Furthermore, it exhibited the lowest RMSE and rRMSE values, indicating greater prediction accuracy. When compared to other models, such as the SVR in both its linear and radial forms, the MLR model using the combined indices consistently outperformed them, as evidenced by its higher R^2 and reduced errors in the RMSE and rRMSE metrics with cross-validation.

Based on the data presented in Table 10, the SVR-rbf model, which utilizes a combination of VIs and TIs, demonstrated the best performance in forecasting grain yield at the flowering and milk stages in 2023. This model reached the highest R^2_CV values at these stages, with an R^2_CV of 0.26 during flowering and 0.24 during the milk stage. Conversely, at the dough stage, the Random Forest model excelled, achieving the highest R^2_CV with a value of 0.20.

Table 10. Assessment of machine learning model performance in forecasting grain yield at various growth stages for 2023, using Leave-One-Out Cross-Validation (CV).

Machine Learning Models	Validation (Leave-One-Out CV)								
	Flowering			Milk			Dough		
	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)	0.11	0.96 (0.63)	10.37 (6.86)	0.12	0.95 (0.62)	10.32 (6.73)	**0.10**	**0.96 (0.59)**	**10.43 (6.37)**
Support Vector Machine (SVR-linear)	**0.16**	**0.93 (0.61)**	**10.07 (6.59)**	**0.15**	**0.94 (0.64)**	**10.13 (6.89)**	−0.06	1.04 (0.70)	11.30 (7.55)
Support Vector Machine (SVR-rbf)	0.14	0.94 (0.53)	10.17 (5.79)	0.11	0.96 (0.63)	10.35 (6.88)	0.00	1.02 (0.63)	11.01 (6.78)
Random Forest (RF)	0.13	0.94 (0.59)	10.23 (6.39)	−0.16	1.09 (0.68)	11.82 (7.35)	−0.19	1.11 (0.68)	11.99 (7.35)
Texture Index (TI)									
Multiple Linear Regression (MLR)	0.06	0.98 (0.58)	10.64 (6.29)	0.05	0.99 (0.65)	10.71 (7.07)	−0.04	1.03 (0.70)	11.19 (7.58)
Support Vector Machine (SVR-linear)	0.06	0.99 (0.62)	10.69 (6.66)	0.12	0.95 (0.59)	10.34 (6.42)	−0.03	1.03 (0.67)	11.18 (7.27)
Support Vector Machine (SVR-rbf)	**0.23**	**0.89 (0.59)**	**9.63 (6.43)**	**0.23**	**0.89 (0.54)**	**9.67 (5.89)**	0.04	0.99 (0.64)	10.78 (6.89)
Random Forest (RF)	−0.06	1.04 (0.60)	11.31 (6.53)	0.18	0.92 (0.51)	9.93 (5.56)	**0.20**	**0.91 (0.59)**	**9.86 (6.36)**
all (IV+TI)									
Multiple Linear Regression (MLR)	0.12	0.95 (0.62)	10.31 (6.74)	0.13	0.95 (0.65)	10.24 (7.04)	0.10	0.96 (0.59)	10.43 (6.37)
Support Vector Machine (SVR-linear)	0.16	0.93 (0.61)	10.07 (6.59)	0.12	0.95 (0.59)	10.34 (6.42)	−0.07	1.05 (0.68)	11.36 (7.33)
Support Vector Machine (SVR-rbf)	**0.26**	**0.87 (0.54)**	**9.45 (5.81)**	**0.24**	**0.88 (0.54)**	**9.56 (5.90)**	0.02	1.01 (0.64)	10.89 (6.91)
Random Forest (RF)	0.11	0.96 (0.58)	10.35 (6.32)	0.04	0.99 (0.62)	10.75 (6.73)	**0.11**	**0.96 (0.59)**	**10.37 (6.40)**

Note: Values in bold represent the best performance achieved by the machine learning models at each phenological stage (Flowering, Milk, and Dough).

Based on the information provided in Table 11, the machine learning models that demonstrated superior effectiveness in forecasting grain yield across various phenological phases for the 2022–2023 period were the RF and MLR models. At the flowering stage, the RF model obtained the leading R^2_CV value of 0.45. Conversely, during the milk stage, the MLR model excelled, achieving the top R^2_CV value of 0.52, surpassing other models. Lastly, in the dough stage, the RF model once again emerged as the leader, attaining the highest R^2_CV value of 0.58. These results underscore the RF model's outstanding performance in the dough stage and the MLR model's superiority in the milk stage.

Table 11. Assessment of machine learning model performance in forecasting grain yield at various growth stages for 2022–2023, using Leave-One-Out Cross-Validation (CV).

Machine Learning Models	Validation (Leave-One-Out CV)								
	Flowering			Milk			Dough		
	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)	0.36	1.50 (0.94)	15.32 (9.64)	**0.47**	**1.36 (0.85)**	**13.92 (8.71)**	0.42	1.43 (0.93)	14.57 (9.55)
Support Vector Machine (SVR-linear)	0.30	1.57 (1.12)	16.07 (11.44)	0.40	1.46 (0.96)	14.90 (9.78)	0.39	1.47 (0.96)	15.00 (9.85)
Support Vector Machine (SVR-rbf)	0.41	1.44 (0.99)	14.70 (10.11)	0.36	1.51 (1.03)	15.39 (10.49)	**0.50**	**1.32 (0.92)**	**13.52 (9.37)**
Random Forest (RF)	**0.45**	**1.39 (0.92)**	**14.20 (9.39)**	0.37	1.49 (0.93)	15.26 (9.51)	0.47	1.36 (0.88)	13.95 (9.00)
Texture Index (TI)									
Multiple Linear Regression (MLR)	0.15	1.73 (1.17)	17.72 (11.99)	0.25	1.62 (1.10)	16.59 (11.27)	0.30	1.57 (1.01)	16.09 (10.35)
Support Vector Machine (SVR-linear)	0.10	1.78 (1.30)	18.20 (13.32)	0.18	1.70 (1.21)	17.34 (12.37)	0.22	1.66 (1.18)	16.95 (12.10)
Support Vector Machine (SVR-rbf)	0.16	1.72 (1.23)	17.60 (12.62)	0.18	1.70 (1.19)	17.37 (12.22)	0.44	1.40 (0.98)	14.31 (9.99)
Random Forest (RF)	**0.25**	**1.62 (1.06)**	**16.58 (10.87)**	**0.28**	**1.60 (1.16)**	**16.34 (11.82)**	**0.45**	**1.40 (0.96)**	**14.29 (9.77)**
all (IV+TI)									
Multiple Linear Regression (MLR)	0.42	1.42 (0.89)	14.55 (9.13)	**0.52**	**1.31 (0.84)**	**13.35 (8.62)**	0.56	1.25 (0.82)	12.73 (8.35)
Support Vector Machine (SVR-linear)	0.28	1.59 (1.13)	16.29 (11.56)	0.48	1.36 (0.90)	13.89 (9.25)	0.49	1.35 (0.84)	13.75 (8.61)
Support Vector Machine (SVR-rbf)	0.40	1.45 (0.99)	14.86 (10.14)	0.35	1.52 (1.03)	15.52 (10.58)	0.57	1.24 (0.84)	12.63 (8.64)
Random Forest (RF)	**0.44**	**1.41 (0.89)**	**14.39 (9.13)**	0.46	1.38 (0.92)	14.10 (9.44)	**0.58**	**1.21 (0.81)**	**12.41 (8.24)**

Note: Values in bold represent the best performance achieved by the machine learning models at each phenological stage (Flowering, Milk, and Dough).

4. Discussion

4.1. Importance of Vegetation and Texture Indices in Rice Yield Prediction

Analysis of the R^2 relationship between rice yield and VIs and TIs (Figure 7) highlights the impact of phenological stages on predictive performance. In Figure 7a, the number

of plots evaluated during 2022 is consistently higher in later phenological stages, which likely contributes to better model stability and accuracy compared to 2023. This aligns with previous findings suggesting that larger sample sizes improve yield predictions [15,24,48]. Additionally, the number of evaluated plots significantly influenced prediction stability, with stages involving a greater number of plots exhibiting improved predictive accuracy. This emphasizes the critical role of sample size in enhancing model reliability. Previous studies have recommended a minimum of 10–20 subplots [48], with successful rice yield predictions using sample sizes ranging from 22 to 122 subplots [9,11,15,18]. These findings underline the importance of representative sample sizes ($\geq$15 subplots) for accurate and reliable model development.

4.2. Temporal Variability of Correlations and Their Impact on Prediction

The Pearson correlation coefficient (r) analysis for 2022, 2023, and the combined 2022–2023 period confirms that the milk and dough phenophases were the most suitable for yield prediction in both years (Table 8), with NDRE and GNDVI being the most reliable spectral indices, although interannual variability slightly reduced correlations in the combined analysis. In 2022, the milk stage showed the highest correlations with NDRE (r = 0.63) and GNDVI (r = 0.61), while in the dough stage, NDRE (r = 0.71) and GNDVI (r = 0.73) emerged as the best predictors. In 2023, correlations at the milk stage were lower but still significant, with NDRE at r = 0.35 and GNDVI at r = 0.44, while in the dough stage, most spectral indices showed no significant relationships with yield. In the combined 2022–2023 analysis, the same indices maintained their relevance, with NDRE (r = 0.57) and GNDVI (r = 0.45) in the milk stage and NDRE (r = 0.56) and GNDVI (r = 0.49) in the dough stage. In contrast, the flowering phenophase was not a reliable stage for yield estimation in either year, showing weak or negative correlations in most indices, such as NDVI and GNDVI in the combined analysis. However, TIs showed more variable and inconsistent correlations with yield in both years, with significant negative values in the milk stage, such as ENTR (r = $-$0.34) and CORR (r = $-$0.32) in the combined period, suggesting that lower texture variability is associated with higher yields. In 2022, some TIs showed negative correlations in the milk stage, such as CONTR (r = $-$0.49) and VAR (r = $-$0.50), while in 2023, correlations were weaker or not significant, indicating that the relationship between canopy texture and yield may depend on environmental and structural factors of the crop. These results highlight the importance of integrating spectral and textural information to improve yield prediction under variable climatic conditions. The higher performance of textural indices in 2023 likely reflects their ability to capture canopy structural variability, which becomes more pronounced under extreme climatic conditions. This underscores the complementary role of TIs alongside VIs in enhancing the accuracy of yield prediction models, particularly in challenging environmental scenarios [25,31].

4.3. Performance of Machine Learning Models in Different Scenarios

The analysis highlights the effectiveness of different prediction models in estimating grain yield across various phenological stages, as detailed in Tables 9–11. In 2022, the MLR model, which integrates VIs and TIs using LOOCV, achieved the highest R^2_CV (0.69) during the flowering phase when combining VIs and TIs (VIs + TIs). In the maturation phase, particularly during the dough stage, which is closer to the harvesting period, the highest R^2_CV recorded was 0.78. Additionally, MLR demonstrated higher accuracy with lower RMSE and rRMSE values, underscoring its robustness and reliability. These results align with those of Duan et al. [17], who reported R^2_CV values ranging from 0.48 to 0.62 using LOOCV, as well as with Wang et al. [24], who demonstrated the effectiveness of integrating VIs and TIs in specific phenological phases, achieving an R^2_CV = 0.86. In

2023, model performance varied due to extreme meteorological conditions. The Support Vector Machine with a radial basis function kernel (SVR-rbf) was the best predictor during the flowering (R^2_CV = 0.26) and milk (R^2_CV = 0.24) phases, whereas the RF model performed best in the dough phase (R^2_CV = 0.20). However, model reliability in 2023 was affected by extreme climatic events, particularly Cyclone Yaku, which significantly disrupted yield prediction in northern coastal Peru [49]. The intense and sudden rainfall associated with this event led to soil saturation, reducing aeration and impairing the absorption of essential nutrients for rice growth, while also disrupting phenological cycles by delaying critical stages such as flowering and grain filling, ultimately reducing productivity. These extreme climatic conditions expose the limitations of conventional models, which struggle to accurately capture climate variability and uncertainty. According to Shourie et al. [50], extreme weather events can significantly alter soil physical properties, affecting its water and nutrient retention capacity, compromising fertility, and potentially leading to salinization, reduced nutrient and water availability, disruptions in carbon and nitrogen cycles, and biodiversity loss, all of which increase crop vulnerability and reduce agricultural yields. Given that rising temperatures, intense rainfall, and soil saturation directly impact agricultural productivity, yield modeling should incorporate holistic approaches that integrate dynamic environmental variables and extreme climatic events. Additionally, studying species highly sensitive to extreme weather conditions provides valuable insights into the interplay between climate, biodiversity, and agricultural sustainability, supporting more resilient and adaptive farming systems [51].

In the combined 2022–2023 period, the RF and MLR models were the most prominent. During the flowering phase, the RF model achieved the highest R^2_CV (0.45), while in the milk phase, the MLR model performed better (R^2_CV = 0.52). In the dough phase, RF regained dominance with R^2_CV = 0.58. These results align with the study by Quille-Mamani et al. [16], who used Sentinel-2 images and phenological metrics based on NDVI to predict rice yield. Although their RF model showed lower accuracy (R^2_CV = 0.44) and a higher RMSE (1.66 t ha^{-1}), these differences could be attributed to variability in input variables and the frequency of satellite data acquisition. Unlike this study, which integrates both VIs and TIs, Quille-Mamani et al. [16] focused exclusively on NDVI-derived metrics, potentially limiting their model's predictive capability. On the other hand, studies such as Wan et al. [52] have demonstrated that incorporating additional variables, such as canopy height and multiple vegetation indices, improves the performance of RF models (R^2 = 0.85; RMSE = 0.39 t ha^{-1}). However, these models remain vulnerable to extreme climatic events, such as Cyclone Yaku and El Niño Costero, which introduce significant variability in agricultural yields. Historical events, such as El Niño 2017, have highlighted the widespread impact of these anomalies on agriculture, the economy, and public health, emphasizing the need to integrate climatic and socioeconomic variables into agricultural planning [53].

These findings underscore the complementarity and adaptability of the RF and MLR models in capturing yield variability across different phenological phases. While MLR performed better under stable conditions, its sensitivity to climatic variations highlights the importance of integrating meteorological variables to enhance its robustness in diverse scenarios. Previous studies have demonstrated that incorporating environmental factors can significantly improve predictive accuracy [6,18], reinforcing the potential of hybrid approaches that combine spectral, textural, and climatic datasets for yield modeling.

4.4. Applicability of the Models and Considerations for Future Prediction

Commercial fields and the most cultivated rice varieties in the Lambayeque region were considered, which allowed the development of models based on real production

conditions with potential for scalability throughout the rice-growing area. However, data collection for yield estimation faced limitations due to the difficulty of obtaining yield data from a sufficient number of subplots, which impacted the accuracy of the predictions. In addition, in 2023, Cyclone Yaku caused significant weather anomalies that altered yield patterns and highlighted the need to integrate meteorological variables into prediction models. Future research should incorporate these variables (precipitation, relative humidity, temperature and others), as their inclusion could improve the accuracy of models and their resilience to environmental fluctuations [15,24]. In this context, extreme events such as Cyclone Yaku underline the importance of developing adaptive and dynamic models capable of capturing extreme climate variability, ensuring the sustainability of agricultural production in regions vulnerable to sudden environmental changes.

5. Conclusions

This study demonstrates that the integration of spectral and textural indices derived from UAV imagery significantly enhances rice yield prediction in the Lambayeque region, Peru. The evaluated machine learning models exhibited variability in performance depending on the phenological stage and the dataset analyzed. In 2022, MLR achieved the best performance during the dough stage, with NDVI, NDRE, and DVI identified as the most relevant spectral indices, while ENT and CON were the most influential textural indices. In 2023, SVR-rbf was the most accurate model during the milk stage, where NDRE and GNDVI were the most important spectral indices, and VAR and COR were the most significant textural indices. The combination of data from both years established RF as the most robust model, producing the best predictions in the milk and dough stages, with NDRE, GNDVI, ENT, and VAR as the key predictive indices. The models were validated using LOOCV, ensuring their stability and minimizing estimation bias, thus providing a robust assessment of predictive capability.

A key factor influencing the effectiveness of predictive models is the rate of fertilizer application, particularly nitrogen, phosphorus, and potassium (NPK). Although the evaluated fields used similar NPK doses (as described in Section 2.4), previous studies have demonstrated that different nitrogen application rates can improve yield prediction accuracy. For instance, Wang et al. [24] found that variations in nitrogen fertilization not only affected crop yield but also influenced the textural characteristics captured in UAV images, suggesting that differentiated fertilizer management could further optimize predictive models.

These findings highlight the importance of integrating spectral, textural, and agronomic data, as well as the necessity of incorporating meteorological and edaphic variables to enhance model generalization across different production cycles. Future research should focus on exploring advanced modeling strategies, such as hybrid approaches and deep learning, while also considering variability in fertilizer management as a key factor in yield prediction. This study provides valuable insights for the implementation of precision agriculture and contributes to more efficient and sustainable crop management.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/rs17040632/s1: Figure S1: Results of sequential feature selection for MLR and SVR models using vegetation indices (VIs), texture indices (TIs), and their combination (VI+TI) at the 2022 flowering stage; Figure S2: Results of sequential feature selection for MLR and SVR models using vegetation indices (VIs), texture indices (TIs), and their combination (VI+TI) at the 2022 milk stage; Figure S3: Results of sequential feature selection for MLR and SVR models using vegetation indices (VIs), texture indices (TIs), and their combination (VI+TI) at the 2022 dough stage; Figure S4: Results of sequential feature selection for MLR and SVR models using vegetation indices (VIs), texture indices (TIs), and their combination (VI+TI) at the 2023 flowering stage; Figure S5: Results of sequential feature selection for MLR

and SVR models using vegetation indices (VIs), texture indices (TIs), and their combination (VI+TI) at the 2023 milk stage; Figure S6: Results of sequential feature selection for MLR and SVR models using vegetation indices (VIs), texture indices (TIs), and their combination (VI+TI) at the 2023 dough stage; Figure S7: Results of sequential feature selection for MLR and SVR models using vegetation indices (VIs), texture indices (TIs), and their combination (VI+TI) at the 2022–2023 flowering stage; Figure S8: Results of sequential feature selection for MLR and SVR models using vegetation indices (VIs), texture indices (TIs), and their combination (VI+TI) at the 2022–2023 milk stage; Figure S9: Results of sequential feature selection for MLR and SVR models using vegetation indices (VIs), texture indices (TIs), and their combination (VI+TI) at the 2022–2023 dough stage; Table S1: Comparative assessment of machine learning models for forecasting crop outcomes at different phenological stages in 2022; Table S2: Comparative assessment of machine learning models for forecasting crop outcomes at different phenological stages in 2023; Table S3: Comparative assessment of machine learning models for forecasting crop outcomes at different phenological stages in 2022–2023; Table S4: Optimization of Random Forest parameters (number of trees and variables) to predict crop yields using different combinations of indices at different phenological stages of 2022; Table S5: Optimization of Random Forest parameters (number of trees and variables) to predict crop yields using different combinations of indices at different phenological stages of 2023; Table S6: Optimization of Random Forest parameters (number of trees and variables) to predict crop yields using different combinations of indices at different phenological stages of 2022–2023.

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

Funding: This research was funded by the National Scientific Research and Advanced Studies Program (PROCIENCIA) of PROCIENCIA-Peru, under the project "Implementation of technological tools in the development of a precision system with remote sensors to optimize the use of water and reduce the emission of greenhouse gases in rice fields for the benefit of farmers in the Lambayeque region" (contract No. PE501078113-2022).

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

Acknowledgments: This paper was written as a result of a collaboration that was initiated by contract No. PE501078113-2022. We also thank Fundo Ventura for the availability of commercial rice yield data.

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

References

1. Xu, T.; Wang, F.; Shi, Z.; Miao, Y. Multi-Scale Monitoring of Rice Aboveground Biomass by Combining Spectral and Textural Information from UAV Hyperspectral Images. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *127*, 103655. [CrossRef]
2. OECD; FAO. *OECD-FAO Agricultural Outlook 2023–2032*; OECD Publishing: Paris, France, 2023; ISBN 9789264619333.
3. Jiang, P.; Zhou, X.; Zhang, L.; Liu, M.; Xiong, H.; Guo, X.; Zhu, Y.; Luo, J.; Chen, L.; Liu, J.; et al. Improving Rice Yield by Promoting Pre-Anathesis Growth in Subtropical Environments. *Agronomy* **2023**, *13*, 820. [CrossRef]
4. FAO. FAO STAT Production Data. Available online: https://www.fao.org/faostat/es/#data/QCL (accessed on 14 March 2024).
5. Ramos-Fernández, L.; Gonzales-Quiquia, M.; Huanuqueño-Murillo, J.; Tito-Quispe, D.; Heros-Aguilar, E.; del Pino, L.F.; Torres-Rua, A. Water Stress Index and Stomatal Conductance under Different Irrigation Regimes with Thermal Sensors in Rice Fields on the Northern Coast of Peru. *Remote Sens.* **2024**, *16*, 796. [CrossRef]
6. 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]
7. Islam, S.M.M.; Gaihre, Y.K.; Islam, M.R.; Islam, A.; Singh, U.; Sander, B.O. Effects of Integrated Plant Nutrition Systems with Fertilizer Deep Placement on Rice Yields and Nitrogen Use Efficiency under Different Irrigation Regimes. *Heliyon* **2023**, *9*, e23110. [CrossRef]

8. Fernandes, A.P.G.; Machado, J.; Fernandes, T.R.; Vasconcelos, M.W.; Carvalho, S.M.P. Water and Nitrogen Fertilization Management in Light of Climate Change: Impacts on Food Security and Product Quality. In *Plant Nutrition and Food Security in the Era of Climate Change*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 147–178.
9. Kurihara, J.; Nagata, T.; Tomiyama, H. Rice Yield Prediction in Different Growth Environments Using Unmanned Aerial Vehicle-Based Hyperspectral Imaging. *Remote Sens.* **2023**, *15*, 2004. [CrossRef]
10. Luo, S.; Jiang, X.; Jiao, W.; Yang, K.; Li, Y.; Fang, S. Remotely Sensed Prediction of Rice Yield at Different Growth Durations Using UAV Multispectral Imagery. *Agriculture* **2022**, *12*, 1447. [CrossRef]
11. Zhou, X.; Zheng, H.B.; Xu, X.Q.; He, J.Y.; Ge, X.K.; Yao, X.; Cheng, T.; Zhu, Y.; Cao, W.X.; Tian, Y.C. Predicting Grain Yield in Rice Using Multi-Temporal Vegetation Indices from UAV-Based Multispectral and Digital Imagery. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 246–255. [CrossRef]
12. Franch, B.; Bautista, A.S.; 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. Saravia, D.; Valqui-Valqui, L.; Salazar, W.; Quille-Mamani, J.; Barboza, E.; Porras-Jorge, R.; Injante, P.; Arbizu, C.I. Yield Prediction of Four Bean (*Phaseolus vulgaris*) Cultivars Using Vegetation Indices Based on Multispectral Images from UAV in an Arid Zone of Peru. *Drones* **2023**, *7*, 325. [CrossRef]
14. Saravia, D.; Salazar, W.; Valqui-Valqui, L.; Quille-Mamani, J.; Porras-Jorge, R.; Corredor, F.-A.; Barboza, E.; Vásquez, H.; Casas Diaz, A.; Arbizu, C. Yield Predictions of Four Hybrids of Maize (*Zea mays*) Using Multispectral Images Obtained from UAV in the Coast of Peru. *Agronomy* **2022**, *12*, 2630. [CrossRef]
15. Su, X.; Wang, J.; Ding, L.; Lu, J.; Zhang, J.; Yao, X.; Cheng, T.; Zhu, Y.; Cao, W.; Tian, Y. Grain Yield Prediction Using Multi-Temporal UAV-Based Multispectral Vegetation Indices and Endmember Abundance in Rice. *Field Crops Res.* **2023**, *299*, 108992. [CrossRef]
16. Quille-Mamani, J.A.; Ruiz, L.A.; Ramos-Fernández, L. Rice Crop Yield Prediction from Sentinel-2 Imagery Using Phenological Metric. *Environ. Sci. Proc.* **2023**, *28*, 16. [CrossRef]
17. Duan, B.; Fang, S.; Gong, Y.; Peng, Y.; Wu, X.; Zhu, R. Remote Estimation of Grain Yield Based on UAV Data in Different Rice Cultivars under Contrasting Climatic Zone. *Field Crops Res.* **2021**, *267*, 108148. [CrossRef]
18. Nakano, H.; Tanaka, R.; Guan, S.; Ohdan, H. Predicting Rice Grain Yield Using Normalized Difference Vegetation Index from UAV and GreenSeeker. *Crop Environ.* **2023**, *2*, 59–65. [CrossRef]
19. Ji, S.; Gu, C.; Xi, X.; Zhang, Z.; Hong, Q.; Huo, Z.; Zhao, H.; Zhang, R.; Li, B.; Tan, C. Quantitative Monitoring of Leaf Area Index in Rice Based on Hyperspectral Feature Bands and Ridge Regression Algorithm. *Remote Sens.* **2022**, *14*, 2777. [CrossRef]
20. Derraz, R.; Muharam, F.M.; Nurulhuda, K.; Jaafar, N.A.; Yap, N.K. Ensemble and Single Algorithm Models to Handle Multicollinearity of UAV Vegetation Indices for Predicting Rice Biomass. *Comput. Electron. Agric.* **2023**, *205*, 107621. [CrossRef]
21. Wu, Y.; Wang, W.; Gu, Y.; Zheng, H.; Yao, X.; Zhu, Y.; Cao, W.; Cheng, T. SPSI: A Novel Composite Index for Estimating Panicle Number in Winter Wheat before Heading from UAV Multispectral Imagery. *Plant Phenomics* **2023**, *5*, 87. [CrossRef]
22. Mia, M.S.; Tanabe, R.; Habibi, L.N.; Hashimoto, N.; Homma, K.; Maki, M.; Matsui, T.; Tanaka, T.S.T. Multimodal Deep Learning for Rice Yield Prediction Using UAV-Based Multispectral Imagery and Weather Data. *Remote Sens.* **2023**, *15*, 2511. [CrossRef]
23. Zheng, H.; Cheng, T.; Zhou, M.; Li, D.; Yao, X.; Tian, Y.; Cao, W.; Zhu, Y. Improved Estimation of Rice Aboveground Biomass Combining Textural and Spectral Analysis of UAV Imagery. *Precis. Agric.* **2019**, *20*, 611–629. [CrossRef]
24. Wang, F.; Yi, Q.; Hu, J.; Xie, L.; Yao, X.; Xu, T.; Zheng, J. Combining Spectral and Textural Information in UAV Hyperspectral Images to Estimate Rice Grain Yield. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *102*, 102397. [CrossRef]
25. Kelsey, K.; Neff, J. Estimates of Aboveground Biomass from Texture Analysis of Landsat Imagery. *Remote Sens.* **2014**, *6*, 6407–6422. [CrossRef]
26. Fu, Y.; Yang, G.; Song, X.; Li, Z.; Xu, X.; Feng, H.; Zhao, C. Improved Estimation of Winter Wheat Aboveground Biomass Using Multiscale Textures Extracted from UAV-Based Digital Images and Hyperspectral Feature Analysis. *Remote Sens.* **2021**, *13*, 581. [CrossRef]
27. Ma, Y.; Ma, L.; Zhang, Q.; Huang, C.; Yi, X.; Chen, X.; Hou, T.; Lv, X.; Zhang, Z. Cotton Yield Estimation Based on Vegetation Indices and Texture Features Derived From RGB Image. *Front. Plant Sci.* **2022**, *13*, 925986. [CrossRef]
28. Yang, K.; Gong, Y.; Fang, S.; Duan, B.; Yuan, N.; Peng, Y.; Wu, X.; Zhu, R. Combining Spectral and Texture Features of UAV Images for the Remote Estimation of Rice LAI throughout the Entire Growing Season. *Remote Sens.* **2021**, *13*, 3001. [CrossRef]
29. Yuan, W.; Meng, Y.; Li, Y.; Ji, Z.; Kong, Q.; Gao, R.; Su, Z. Research on Rice Leaf Area Index Estimation Based on Fusion of Texture and Spectral Information. *Comput. Electron. Agric.* **2023**, *211*, 108016. [CrossRef]
30. Karmakar, P.; Teng, S.W.; Murshed, M.; Pang, S.; Li, Y.; Lin, H. Crop Monitoring by Multimodal Remote Sensing: A Review. *Remote Sens. Appl.* **2024**, *33*, 101093. [CrossRef]
31. Mishra, P.; Nordon, A.; Asaari, M.S.M.; Lian, G.; Redfern, S. Fusing Spectral and Textural Information in Near-Infrared Hyperspectral Imaging to Improve Green Tea Classification Modelling. *J. Food Eng.* **2019**, *249*, 40–47. [CrossRef]
32. Dilmurat, K.; Sagan, V.; Moose, S. AI-Driven Maize Yield Forecasting Using Unmanned Aerial Vehicle-Based Hyperspectral and Lidar Data Fusion. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2022**, *3*, 193–199. [CrossRef]

33. 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]
34. Shawon, S.M.; Ema, F.B.; Mahi, A.K.; Niha, F.L.; Zubair, H.T. Crop Yield Prediction Using Machine Learning: An Extensive and Systematic Literature Review. *Smart Agric. Technol.* **2025**, *10*, 100718. [CrossRef]
35. Yamaguchi, T.; Sasano, K.; Katsura, K. Improving Efficiency of Ground-Truth Data Collection for UAV-Based Rice Growth Estimation Models: Investigating the Effect of Sampling Size on Model Accuracy. *Plant Prod. Sci.* **2024**, *27*, 1–13. [CrossRef]
36. 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]
37. Porras-Jorge, R.; Ramos-Fernández, L.; Ojeda-Bustamante, W.; Ontiveros-Capurata, R. Performance Assessment of the AquaCrop Model to Estimate Rice Yields under Alternate Wetting and Drying Irrigation in the Coast of Peru. *Sci. Agropecu.* **2020**, *11*, 309–321. [CrossRef]
38. Biswal, S.; Pathak, N.; Chatterjee, C.; Mailapalli, D.R. Estimation of Aboveground Biomass from Spectral and Textural Characteristics of Paddy Crop Using UAV-Multispectral Images and Machine Learning Techniques. *Geocarto Int.* **2024**, *39*, 2364725. [CrossRef]
39. Haralick, R.M.; Dinstein, I.; Shanmugam, K. Textural Features for Image Classification. *IEEE Trans. Syst. Man. Cybern.* **1973**, *SMC-3*, 610–621. [CrossRef]
40. Antoniol, G.; Basco, C.; Ceccarelli, M.; Markus, M.; Moritz, L. r.texture—Generate Images with Textural Features From a Raster Map. Available online: https://grass.osgeo.org/grass-stable/manuals/r.texture.html (accessed on 10 October 2024).
41. Moreno-García, B.; Guillén, M.; Casterad, M.A.; Quílez, D. Uso de imágenes aéreas multiespectrales para estimación del rendimiento en cultivo de arroz. Teledetección: Sistemas Operacionales de Observación de la Tierra, *Actas del XV Congreso de la Asociación Española de Teledetección*, Madrid. 2013, pp. 357–560. Available online: http://www.aet.org.es/?q=congresos (accessed on 20 November 2024).
42. Balabin, R.M.; Lomakina, E.I. Support Vector Machine Regression (SVR/LS-SVM)—An Alternative to Neural Networks (ANN) for Analytical Chemistry? Comparison of Nonlinear Methods on near Infrared (NIR) Spectroscopy Data. *Analyst* **2011**, *136*, 1703–1712. [CrossRef]
43. Raschka, S. MLxtend: Providing Machine Learning and Data Science Utilities and Extensions to Python's Scientific Computing Stack. *J. Open Source Softw.* **2018**, *3*, 638. [CrossRef]
44. Zhang, C.; Xue, Y. Estimation of Biochemical Pigment Content in Poplar Leaves Using Proximal Multispectral Imaging and Regression Modeling Combined with Feature Selection. *Sensors* **2023**, *24*, 217. [CrossRef]
45. Raschka, S.; Mirjalili, V. *Python Machine Learning*, 3rd ed.; Packt Publishing: Birmingham, UK, 2019; Volume 1, ISBN 9781789955750.
46. Fawagreh, K.; Gaber, M.M.; Elyan, E. Random Forests: From Early Developments to Recent Advancements. *Syst. Sci. Control Eng.* **2014**, *2*, 602–609. [CrossRef]
47. Moisen, G. Classification and Regression Trees. In *Encyclopedia of Ecology*; Elsevier: Kidlington, OX, USA, 2008; Volume 1.
48. Keith, T.Z. *Multiple Regression and Beyond*, 2nd ed.; Routledge: New York, NY, USA, 2015.
49. MIDAGRI. Impact of Cyclone Yaku on the Agriculture and Livestock Sector. Available online: https://repositorio.midagri.gob.pe/ (accessed on 14 October 2023).
50. Shourie, A.; Singh, A. Impact of Climate Change on Soil Fertility. In *Climate Change and the Microbiome. Soil Biology*; Springer: Cham, Switzerland, 2021; Volume 63, pp. 49–62. [CrossRef]
51. Warner, J.; Alaica, A.K. Contextualizing the Influence of Climate and Culture on Bivalve Populations: Donax Obesulus Malacology from the North Coast of Peru. *J. Isl. Coast. Archaeol.* **2024**, *19*, 150–171. [CrossRef]
52. Wan, L.; Cen, H.; Zhu, J.; Zhang, J.; Zhu, Y.; Sun, D.; Du, X.; Zhai, L.; Weng, H.; Li, Y.; et al. Grain Yield Prediction of Rice Using Multi-Temporal UAV-Based RGB and Multispectral Images and Model Transfer—A Case Study of Small Farmlands in the South of China. *Agric. For. Meteorol.* **2020**, *291*, 108096. [CrossRef]
53. Yglesias-González, M.; Valdés-Velásquez, A.; Hartinger, S.M.; Takahashi, K.; Salvatierra, G.; Velarde, R.; Contreras, A.; María, H.S.; Romanello, M.; Paz-Soldán, V.; et al. Reflections on the Impact and Response to the Peruvian 2017 Coastal El Niño Event: Looking to the Past to Prepare for the Future. *PLoS ONE* **2023**, *18*, e0290767. [CrossRef] [PubMed]