remote sensing

MDPI

Article

Maize Yield Prediction via Data Fusion of UAV Multi/Hyperspectral Imagery and In-Field Measurements

Claudia Savarese [1], Marco De Mizio [1], Francesco Tufano [1], Davide Savy [2], Vincenzo Di Meo [2], Massimiliano Gargiulo [1,*], Sara Parrilli [1] and Vincenza Cozzolino [2]

1 Italian Aerospace Research Centre, Via Maiorise snc, 81043 Capua, Italy; c.savarese@cira.it (C.S.); m.demizio@cira.it (M.D.M.); f.tufano@cira.it (F.T.); s.parrilli@cira.it (S.P.)

2 Department of Agricultural Sciences, University of Naples Federico II, Piazza Carlo di Borbone 1, 80055 Portici, Italy; davide.savy@unina.it (D.S.); vincenzo.dimeo@unina.it (V.D.M.); vincenza.cozzolino@unina.it (V.C.)

* Correspondence: m.gargiulo@cira.it; Tel.: +39-0823-62-31-64

Highlights

What are the main findings?
- Maize grain yield prediction strongly depended on the crop phenological stage: early vegetative prediction benefited from data fusion between UAV-derived spectral indices and in-field measurements, whereas pre-harvest prediction was mainly driven by red-edge vegetation indices and LAI. Yield prediction accuracy strongly depends on phenological stage.
- Multispectral and hyperspectral UAV data showed comparable predictive performance for yield estimation, while validation strategies dramatically affect perceived accuracy.

What is the implication of the main finding?
- For maize yield prediction under similar conditions, multispectral sensors equipped with red-edge bands represent a cost-effective and operationally simpler alternative to more expensive hyperspectral systems.
- To avoid overestimating model generalization beyond the experimental design, validation strategies should explicitly account for the experimental design (e.g., using LOTO or spatial cross-validation), especially when working with small agronomic datasets structured by treatment.

Abstract

Timely forecasting of maize productivity is essential to support precision agriculture and optimize management practices. In this study, we analyzed the potential of integrating ground-based measurements and UAV-derived spectral data for predicting maize grain yield (GY) under different fertilization conditions. Field data were collected at two key phenological stages: early vegetative stage (V7) and pre-harvest (R4). Ground-based measurements included SPAD, above-ground biomass (AGB), and leaf area index (LAI), while multispectral and hyperspectral imagery was acquired by drone. A series of Ordinary Least Squares (OLS) models was developed to evaluate the predictive performance of individual variables and their combinations. Model robustness was assessed using two validation strategies: Leave-One-Treatment-Out (LOTO) to assess model performance across the treatments included in the experimental design and random sampling to assess performance within the dataset. The results showed that yield prediction was less accurate during the early growth stages, where data fusion significantly improved the model's accuracy (R^2 = 0.82; MAE = 6.36 q ha^{-1}; $MAPE \approx$7 %). The predictive performance of VIs alone increased substantially in the pre-harvest stage, with the combination of red-edge indices and LAI proving to be the best model for late yield prediction (R^2 = 0.86;

Check for updates

Academic Editors: Xiaoyang Zhang, Pedro Marques and Leilson Ferreira Gomes

Received: 6 May 2026
Revised: 21 July 2026
Accepted: 22 July 2026
Published: 27 July 2026

Copyright: © 2026 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.

MAE = 6.56 q ha^{-1}; $MAPE$ ≈7%). Comparison of multispectral and hyperspectral data revealed comparable predictive performance, suggesting that multispectral sensors may already capture the key spectral information needed for yield forecasting. Furthermore, random validation consistently produced more optimistic results than the LOTO method, highlighting the importance of using validation strategies that explicitly account for the experimental design when evaluating model performance across the treatments included in the study. Overall, the present study demonstrates that yield prediction is highly dependent on the phenological stage and validation approach, and that integrating complementary data sources can improve model performance, particularly during the early growth stages. These findings should be interpreted as a proof-of-concept based on a single-site, single-season experiment with a limited sample size (n = 12), and therefore require further validation across multiple environments and growing seasons.

Keywords: precision agriculture; UAV multispectral/hyperspectral imaging; proximal sensing; vegetation indices; maize yield prediction; data fusion

1. Introduction

The estimated growth of the global population from seven to nine billion people by 2050 has raised concerns about the need to increase global food production by 60–110% [1]. Against this backdrop, conventional farming systems have intensified the use of chemical fertilizers and pesticides to boost crop yields, despite the well-known harmful consequences for both ecosystem and human health (i.e., soil degradation, greenhouse gas emissions, environmental pollution). These drawbacks, combined with rising consumer interest in environmental sustainability and food safety, have driven modern agriculture to implement farming practices that support the health of soil, ecosystems, and people. In this context, organic farming has gained increasing global popularity as a production system that avoids the use of artificial inputs, relying instead on conservative management practices and/or fertilizers of organic origin, such as compost, manure, and natural waste [2]. However, organic agriculture still faces significant challenges, including the so-called “yield gap” with conventional farming, the difficulties in nutrient and disease management, and the need for specialized knowledge [3]. In particular, organic inputs are often highly variable depending on factors such as raw material composition or composting processes, making their effects on crop growth difficult to predict. In this context, the development of reliable tools for monitoring and predicting crop performance becomes essential to support farmers in making timely and informed management decisions.

Precision agriculture (PA) technologies represent an effective and efficient strategy for the implementation of agricultural management systems to increase both productivity and sustainability [4]. PA methods are based on the use of different sensors for proximal and remote detection of key plant parameters such as chlorophyll content, leaf cover, or nutritional/water status, which are strictly related to crop vigor and yield. All collected information is analyzed alongside fundamental data (i.e., weather and soil conditions) to provide useful indications to farmers, such as the quantity of fertilizer to apply, the right time to intervene, or specific portions of the field to target, thus ensuring sustainable profitability [5]. The ability to collect large volumes of information from multiple sources and different parts of the field through non-invasive and fast measurements is hence essential in precision agriculture. Consequently, remote sensing (RS) techniques are gradually replacing traditional ones in PA applications as they allow mapping of large areas in shorter periods and at lower analysis costs [6]. Among RS platforms, unmanned aerial vehicles

(UAVs) equipped with multispectral and hyperspectral sensors are becoming increasingly relevant in agricultural studies, due to the high spatiotemporal resolution required for field monitoring of crop health and productivity [5]. From UAV imaging, a series of vegetation indices (VIs) can be calculated by combining reflectance measurements in several spectral bands influenced by plant chemical and morphological characteristics, thus providing helpful information for effective agricultural management.

Nowadays, VIs such as normalized difference vegetation index (NDVI), green NDVI (GNDVI), and soil-adjusted vegetation indices (EVI, SAVI) have been employed for multiple PA applications, from water and nutrient control to crop yield estimation, as they are sensitive to several agronomic aspects [6,7]. Due to their global importance, cereals are certainly the main research targets in these types of studies. Among agricultural management practices, the estimation of grain yield before harvest is particularly important for the production of low-impact cereals, and its prediction using UAV-collected data is achieving excellent results [8]. A popular method for yield forecasting involves inserting VI features into Machine Learning (ML) models that employ statistical analysis (e.g., regression, classification, clustering) to detect patterns and correlations between data. Most of these studies use machine learning regression modeling, a supervised approach in which model building is based on one or more input variables including ground-measured agronomic parameters, proximal/remote spectral data, or both. The model learns the relationship between these variables to make future predictions of crop yield, thus helping farmers make decisions based on available information. Unlike previous studies, this work does not propose a novel algorithm but rather provides a controlled experimental framework to disentangle the role of phenology, data fusion, and validation strategy on yield prediction accuracy. Specifically, we systematically compare early- versus late-stage predictions under contrasting fertilization treatments, explicitly evaluate validation strategy bias (Leave-One-Treatment-Out vs. random cross-validation), and quantify the benefits of integrating ground-based measurements with UAV-derived spectral data across different phenological phases. Although previous studies have used phenological information for yield prediction, the novelty of our work lies in the explicit quantification of prediction performance as a function of growth stage under different validation strategies (LOTO vs. random), and in the systematic assessment of whether data fusion provides added value at each stage. Using this approach, many authors have been able to estimate cereal yields with high accuracy, even under different agricultural management practices (e.g., water and nutritional regimes) [9–11]. For example, in [12] multi-temporal hyperspectral imagery and regression models is used to estimate maize yield and flowering time, obtaining a correlation of r = 0.54 between spectral reflectance and yield (RMSE = 2.68 t/ha), and values above r = 0.90 for flowering prediction. Similarly, in [10] it is demonstrated that, even using simple RGB imagery and cumulative vegetation indices (ΣVI-SUMs), cereal production can be estimated with r above 0.65 and an average error of approximately 2 t/ha, while reducing the number of flights required. The effectiveness of this methodology using high-resolution multispectral imagery and machine learning models (Random Forest and Linear Regression) for corn yield forecasting is also confirmed in [13]. In this study, indices based on NIR and Red-Edge bands showed the strongest correlations and lowest estimation errors. The authors also evaluated different spatial validation approaches, underlining the importance of considering the spatial distribution of data to avoid overestimation of accuracy. As stated in [8], a major challenge in crop yield prediction lies in the limited transferability of models across different conditions. Indeed, the performance of ML approaches is strongly influenced by several factors, including crop type, environmental conditions, and management practices, as well as by the limited availability of large and representative datasets. In addition, many existing studies rely on single-date observations or exclusively on remotely sensed data, which may not fully capture the temporal dynamics of crop development or the underlying physiological processes driving yield formation. As a result, grain yield estimation remains a complex task that requires the integration of multiple sources of information. In this context, accounting for crop phenology and combining proximal measurements with multi-source remote sensing data may represent a promising strategy to improve the robustness and accuracy of yield prediction models in precision agriculture applications.

The present study aims to evaluate a phenology-aware data fusion approach for the estimation of maize grain yield under contrasting fertilization practices (mineral vs. organic). Specifically, we investigate how the integration of UAV-based multispectral and hyperspectral imagery with proximal measurements influences prediction performance across different crop growth stages. To this end, data were collected at two key phenological phases: the vegetative stage, when UAV multispectral imagery, SPAD readings, and destructive above-ground biomass (AGB) measurements were acquired, and the pre-harvest stage, when UAV multispectral and hyperspectral imagery were collected alongside SPAD and LAI measurements. A set of ordinary least squares (OLS) linear regression models was developed by varying the combination of input features (proximal and remotely sensed variables) to systematically assess their individual and combined predictive power. This framework enables the identification of the most effective feature configurations for yield estimation at different corn growth stages. A preliminary analysis of part of this dataset was previously presented in [14], where simple correlations between ground measurements and vegetation indices were explored. In the present study, we extend that initial work by implementing a structured predictive modeling approach and explicitly investigating the contribution of multi-temporal and multi-source data integration to grain yield estimation.

2. Materials and Methods

2.1. Study Area and Experimental Setup

The present study was conducted in 2023 at the University of Naples Federico II experimental farm at Castel Volturno (CE) (41°03′47″N, 13°59′32″E), shown in Figure 1. The soil of the study area is classified as Vitrandic Haplucept (USDA soil taxonomy), developed on ash and pumice pyroclastic parent material, and characterized by a clayey texture. The study area is classified as having a Mediterranean temperate climate, with a mean annual temperatures of 18.5 °C and mean annual precipitation of approximately 653 mm, summer being the driest period and autumn being the rainiest period.

The design of the field experiment was directed by the project managers of the Green Farm project (CUP B66G20000660005). The experimental setup consisted of a randomized complete block design with four treatments per block and three replications. The plot size was 40 m^2 (8×5 m). The site was sown with maize (*Zea mays* L.) in early summer (10 June 2023) in rows separated by 0.75 m; maize was spaced 0.25 m within rows, resulting in a plant population density of 7 plants m^2. During the whole maize growing cycle, water was uniformly applied with a sprinkler irrigation system, according to the crop water requirements. The treatments consisted of supplying the same quantity of nitrogen but in different forms (mineral versus organic) at sowing. A graphical view of the experimental design is shown in Figure 1.

In July (26 July 2023), 16 plants per treatment were manually harvested in the central rows of each plot, and the above-ground biomass (AGB, g) was determined by drying the plants (including leaves and stems) in an oven at 70 °C for 72 h and weighing them on an analytical balance. In October (2 October 2023), mechanical harvesting was performed, and grain yield (GY) was determined for each experimental plot. Grain weight was recorded and yield was expressed in q ha^{-1} after adjustment to a standard moisture content (14%).

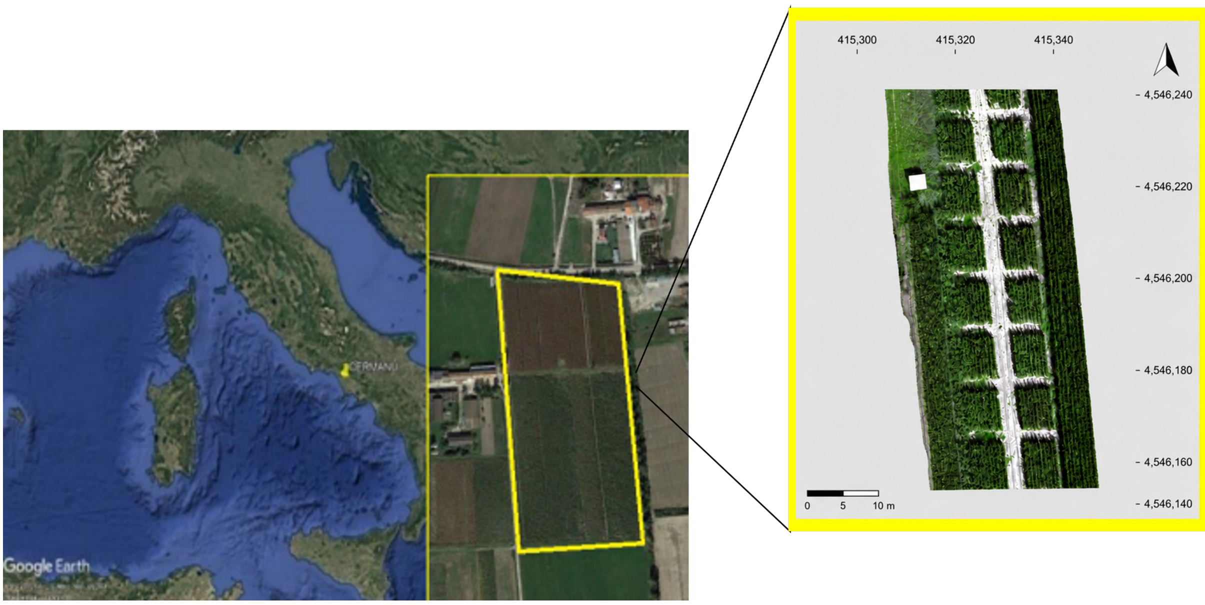

Figure 1. Overview of the study area and experimental design, including location, field layout, and treatment plots. Map coordinates are expressed in the projected UTM reference system of the georeferenced orthomosaic, in meters.

2.2. Ground-Optical Measurements

Readings with two different optical sensors were taken at ground level on two different dates. TTo detect early effects of the four fertilization practices on maize plants, both in terms of development and physiology, the first sampling date (26 July 2023) occurred 40 days after sowing (DAS), when the crop was close to the growth stage referred to as "stem elongation" (about seven fully unfolded leaves, V7). The second sampling date (27 September 2023) was carried out shortly before harvest, when the crop was in the ear maturation stage (R4) and the differences between agronomic treatments were expected to be most evident.

The MultispeQ handheld device (PhotosynQ Inc., East Lansing, MI, USA) was used to measure the Soil Plant Analysis Development (SPAD) index on both sampling dates, while the Leaf Area Index (LAI) was recorded only before harvesting using the LaiPen LP110 (Photon Systems Instruments sro., Drásov, Czech Republic). Both instruments are described in [14]. In particular, SPAD measurements were taken from the uppermost fully developed leaf of 5 representative plants in the two central rows of three plots per treatment. Measurements were taken on the upper side at the longitudinal center of the leaf. The representative value of each treatment was obtained as the average of the 15 readings. On the other hand, LAI measurements were carried out at a medium distance between the maize culms and at about 50 cm from the ground, following a line parallel to the three central rows of three plots for each treatment. All measurements were performed between 10:00 a.m. and noon on sunny days.

Unfortunately, the LaiPen device was not available during the first field survey (July), so dry biomass values were used as structural indicators of corn growth. On the other hand, SPAD and LAI measurements were conducted simultaneously during the second campaign (September), to obtain information on both the health and development of maize plants.

2.3. UAV Imagery

During field surveys, two flights with a customized Foxtech Hover 1 quadcopter were carried out over the maize field. All acquisitions were conducted under clear skies between 10:00 a.m. and noon to minimize BRDF and shadowing effects. The selected flight altitude was around 20 m on both dates, and the spatial resolution of the MS and HS images was 1 cm and 8 cm, respectively. For both sensors, reflectance data were generated using calibrated Reflectance Panels, enabling more accurate compensation for incident light conditions. The difference between these two configurations lies in the data acquisition approach. The first maximizes spatial resolution in the VIS-NIR spectrum, sacrificing spectral resolution, whereas the second has exactly the opposite approach.

At the early-vegetative survey (26 July 2023), the drone was equipped with the Micasense Altum (MicaSense Inc., Wichita, KS, USA) multispectral (MS) sensor, which has six spectral bands, including blue, green, red, red-edge, near-infrared, and thermal. Thanks to six different Lens/Sensor pairs, six distinct images are produced at different sizes. The resolution of images related to the bands operating in the VIS-NIR spectrum is 2064×1544 pixels. Additionally, the MicaSense Altum is integrated with a GPS and a Downwelling Light Sensor (DLS) to measure ambient light and sun angle, which is crucial for correcting global lighting changes during flights.

At the pre-harvest survey (28 September 2023), the Cubert Ultris 5 (Cubert GmbH, Ulm, Germany) hyperspectral (HS) sensor was added to the MS setup. This camera provided 51 bands covering the visible and near-infrared ranges and acquired images with a global shutter sensor rather than a push-broom sensor. The image resolution was 290×275 pixels for each band.

Technical specifications of the sensors are reported in Table 1. Radiometric calibration followed manufacturer-recommended procedures. Multispectral (MicaSense Altum) imagery was corrected to reflectance using calibrated reference panels acquired immediately before/after each flight and the DLS to compensate short-term irradiance fluctuations; hyperspectral (Cubert ULTRIS 5) cubes were converted to reflectance with the vendor workflow and then processed identically.

Table 1. Technical specifications comparison: MicaSense Altum vs. Cubert Ultris 5.

Specification	**MicaSense Altum**	**Cubert Ultris 5**
Sensor Type	Multispectral + Thermal	Hyperspectral
Acquisition Technology	5 high-resolution lenses + 1 LWIR microbolometer (global shutters on VNIR)	2D Snapshot (Global Shutter)
Spectral Bands	6 discrete bands: - Blue (475 nm, 20 nm BW) - Green (560 nm, 20 nm BW) - Red (668 nm, 10 nm BW) - Red Edge (717 nm, 10 nm BW) - NIR (840 nm, 40 nm BW) - Thermal (8–14 µm)	51 contiguous bands: - Range: 450–850 nm (VIS-NIR) - Sampling: 8 nm - FWHM: 26 nm @ 532 nm
Spatial Resolution	Multispectral: 2048×1536 px Thermal: 160×120 px	290×275 px
Ground Sampling Distance (GSD) at 120 m	Multispectral: 4.0 cm/px Thermal: 64 cm/px	7.2 cm/px
Field of View (FOV)	Multispectral: 49×38 Thermal: 57×44	15° (horizontal)
Focal Length	Multispectral: 8.0 mm	17 mm (equivalent)
Bit Depth	16 bit	12 bit
Maximum Frame Rate	1 Hz (typical for mapping)	15 Hz
Weight	470 g (camera only)	126 g
Data Output	Co-registered, discrete band images	Hyperspectral Data Cube
Integrated Calibration	Yes (DLS 2 & Calibration Panel)	Requires separate calibration panel

2.4. Image and Data Processing

Agisoft Metashape software (version 2.2.2, https://www.agisoft.com/) was used to generate orthomosaics from the reflectance MS and HS images, yielding ground sample distances of ≈1 cm (MS) and ≈8 cm (HS), after the calibration phase. MATLAB software (version R2021a) was then employed to compute a set of multispectral and hyperspectral indices commonly used in precision agriculture applications. Plot-level spectra were extracted from fixed, centrally placed rectangular ROIs to buffer plot edges and inter-row paths; the July mosaics were additionally masked with NDVI > 0.3 to suppress soil background, whereas the September scenes (near full cover) were processed without thresholding. The list of all calculated Vegetation Indices (VIs) is reported in the Supplementary Materials (Tables S1 and S2).

2.5. Vegetation Index Selection

In order to evaluate the relationships among the acquired variables, Pearson's correlation coefficients (r) were computed between vegetation indices (VIs), agronomic parameters (SPAD, AGB, and LAI), and grain yield (GY). The Pearson coefficient is defined as:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \ \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

where x_i and y_i are the n observations of the two variables, while $\bar{x}$ and $\bar{y}$ are their respective means.

Correlation analysis was performed to (i) identify vegetation indices most strongly related to grain yield and (ii) assess potential redundancy between remotely sensed indices and ground-based measurements.

Data acquisition differed by date and crop stage:

- in July (vegetative stage) UAV multispectral (MS) imagery was acquired together with SPAD measurements and destructive above-ground biomass (AGB);
- in September (pre-harvest) UAV multispectral (MS) and hyperspectral (HS) imagery was acquired together with SPAD and leaf area index (LAI) measurements.

Vegetation indices, reported in Table 2, were then selected separately for each acquisition date based on their correlation with grain yield and their complementarity with ground measurements. Due to the limited sample size, correlation results were interpreted in combination with physiological relevance rather than statistical significance alone.

Table 2. Vegetation indices (VIs) derived from multispectral (MS) and hyperspectral (HS) data and selected as predictors for yield estimation.

Vegetation Index	Acronym	Formula
MS and HS data		
Normalized Difference Vegetation Index	NDVI	$(NIR - R)/(NIR + R)$
Green NDVI	GNDVI	$(NIR - G)/(NIR + G)$
Red-Edge NDVI	NDRE	$(NIR - RE)/(NIR + RE)$
Leaf Chlorophyll Index	LCI	$(NIR - RE)/(NIR + R)$
Structure Intensive Pigment Index	SIPI	$(NIR - B)/(NIR - R)$
Modified Chlorophyll Absorption in Reflectance Index	mCARI	$1.2(2.5(NIR - R) - 1.3(NIR - G))$
Transformed Vegetation Index	TVI	$0.5(120(RE - G) - 200(R - G))$
Only HS data		
Gitelson and Merzlyak 2	GM_2	R_{750}/R_{700}
Vogelmann Red Edge Index 1	VOG_1	R_{740}/R_{720}

Finally, selected predictors were used as inputs to regression models for early (vegetative) and late (pre-harvest) grain yield estimation.

2.6. Regression Techniques for Yield Estimation

To evaluate the predictive power of ground-based variables and vegetation indices derived from UAV data, ordinary least squares (OLS) regression models were developed to estimate maize grain yield (GY).

Several sets of predictors were considered, including:

- Ground-measured agronomic variables (SPAD, AGB or LAI, depending on the survey);
- VIs derived from UAV-acquired multispectral and hyperspectral data;
- Data-fusion between field data and VIs.

In particular, the following models were tested for each vegetation index:

- VI only (best single VI per dataset);
- VI combined with SPAD;
- VI combined with the second ground-based variable (AGB or LAI);
- VI combined with both SPAD and a structural ground-based variable.

Models based exclusively on agronomic variables (SPAD, AGB/LAI and their combination) and models that simultaneously include all available indices were also evaluated. Ordinary least squares (OLS) linear regression models were preferred over more complex machine learning approaches (e.g., Random Forest, SVM) due to the extremely limited sample size (n = 12 experimental plots). With such a small dataset, more flexible models would risk severe overfitting and poor generalizability. OLS provides a transparent, interpretable baseline and allows direct assessment of individual predictor contributions. All analyses were implemented in the R programming language, using linear regression (lm) models and custom functions for cross-validation and metric calculation.

2.6.1. Validation Strategies

Given the limited but experimentally structured dataset, consisting of 12 samples from four treatments with three replicates each, the linear regression model was tested and evaluated using multiple complementary validation schemes, described in detail below, rather than relying on a single split between training and testing sets. This choice was made not only to account for the high sensitivity of performance estimates in regression settings with small samples and to obtain a more conservative assessment of the model's behavior but also to evaluate the validation strategy under different a priori knowledge conditions. The different validation strategies were therefore designed to assess both the models' predictive power and robustness with respect to the experimental conditions.

Leave-One-Treatment-Out (LOTO)

The primary validation was conducted using the Leave-One-Treatment-Out (LOTO) approach. In this design, at each iteration, an entire treatment (A, B, C, or D) is excluded from the training set and used as the test set, while the model is trained on data from the remaining three treatments.

This procedure allows us to conduct a 4-fold analysis and to evaluate the model's generalization capability in terms of agronomic conditions not observed during training.

Performance metrics were calculated for each fold, and the final results were obtained as the average of the values across the four folds.

Random Cross-Validation

A random cross-validation was also implemented, in which at each iteration, nine samples were randomly selected for training and the remaining three for testing. The procedure was repeated 100 times, considering each time a random extraction of the 3 samples for testing (without ever repeating the same triplets), to obtain a robust estimate of the model's average performance.

This strategy allows us to evaluate the average predictive capability of the model but does not explicitly take into account the experimental design.

2.6.2. Evaluation Metrics

The performance of the models was evaluated using the coefficient of determination (R^2), the Mean Absolute Error (MAE), and the Mean Absolute Percentage Error (MAPE), defined as follows:

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

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2}$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{3}$$

where y_i represents the observed value, $\hat{y}_i$ the predicted value, $\bar{y}$ the mean of observed values, and n the number of observations.

R^2 provides a measure of the model's ability to explain yield variability, while MAE and MAPE quantify prediction accuracy in absolute and relative terms. Given the primary objective of yield estimation, particular attention was paid to error metrics (MAE and MAPE) in model selection. For clarity, only the average values of R^2, MAE, and MAPE are reported in the main text, while the corresponding standard deviations across validation folds (LOTO) and iterations (Random CV) are provided in the Supplementary Materials. To assess the relative importance of predictors in the best-performing OLS models, standardized regression coefficients (β weights) were calculated. Predictor variables were standardized to zero mean and unit variance prior to model fitting, allowing direct comparison of their relative contribution to grain yield estimation. The statistical significance of each predictor was assessed using the associated p-values.

3. Results

To ensure clarity and avoid redundancy, only the most relevant results are presented in this section, while the complete set of analyses is provided in the Supplementary Materials.

3.1. Correlation Analysis

The Pearson correlation analysis between maize grain yield (GY), ground-based variables, and UAV-derived vegetation indices (VIs) from different datasets is shown in Table 3. Biomass and yield data were provided by Prof. Vincenza Cozzolino and her team of University of Naples Federico II - Project Green Farm (CUP B66G20000660005).

Agronomic variables (SPAD, AGB, and LAI) showed strong positive correlations with GY for both the early and late field surveys.

In the early vegetative stage (V7), only moderate correlations were observed for the UAV-derived multispectral indices. Among the VIs, SIPI showed the strongest positive correlation with yield (r = 0.647, p = 0.02), while the remaining indices presented weaker and non-significant relations ($|r| \approx$ 0.2–0.4). Contrarily, in the pre-harvest phase (R4) correlations were significantly stronger. Multispectral indices such as NDRE (r = 0.849, $p < 0.01$) and NDVI (r = 0.843, $p < 0.01$) showed high positive correlations with yield, while SIPI and TVI displayed strong negative correlations (r = −0.865 and r = −0.838, respectively; $p < 0.01$). Hyperspectral indices (pre-harvest only) further confirmed this trend, with VOG1 showing the strongest positive correlation (r = 0.780, $p < 0.01$), followed by GM2 (r = 0.655, $p < 0.01$), while negative relationships with corn yield were again observed for TVI and SIPI.

Table 3. Pearson correlation coefficients (r) and *p*-values between maize grain yield (GY), ground-based variables, and the selected UAV-derived vegetation indices (VIs) for the early vegetative (V7) and pre-harvest (R4) surveys.

Variable	Early-Vegetative		Pre-Harvest	
	r (GY)	*p*-Value	r (GY)	*p*-Value
Ground-based variables				
SPAD	0.885	<0.001	0.843	<0.01
AGB	0.708	<0.01	–	–
LAI	–	–	0.947	<0.001
Multispectral indices				
SIPI	0.647	<0.05	−0.865	<0.001
LCI	0.400	0.198	0.859	<0.001
NDRE	0.328	0.299	0.849	<0.001
NDVI	0.112	0.728	0.843	<0.01
TVI	−0.324	0.304	−0.838	<0.01
GNDVI	0.287	0.365	0.828	<0.01
mCARI	−0.209	0.514	0.460	0.132
Hyperspectral indices (Pre-harvest only)				
VOG1	–	–	0.780	0.003
NDRE	–	–	0.779	0.003
TVI	–	–	−0.633	0.027
SIPI	–	–	−0.506	0.093
mCARI	–	–	−0.100	0.758

In general, all identified relationships showed approximately linear behavior, as shown in Figures 2 and 3. However, clear differences between maize growth stages can be observed. In the early-vegetative campaign, the relationships appear more scattered, with a higher dispersion of data points around the regression lines for both field (Figure 2) and UAV surveys (Figure 3). In contrast, at the pre-harvest stage, the same relationships become more clearly defined and less dispersed, indicating a stronger and more consistent association of all variables with grain yield. Finally, the correlation analysis underlined some differences between the two UAV datasets acquired pre-harvest. Particularly, the multispectral (MS) data showed slightly more defined and less dispersed relationships than hyperspectral (HS) indices, which exhibited greater variability around the regression line (Figure 3).

Overall, the correlation analysis highlighted a strong influence of maize phenological phase on the relationships between grain yield and both ground- and UAV-based variables.

3.2. Regression Analysis

The performance of OLS models for grain yield (GY) estimation at different maize growth stages is reported in Tables 4–6.

In the early vegetative stage (July, Table 4), models based on agronomic variables showed moderate performance using LOTO validation, with R^2 values ranging from 0.455 (AGB) to 0.555 (SPAD), and an improvement in terms of error for the combined $SPAD + AGB$ model (MAE = 8.24 q ha^{-1}; MAPE = 9.54%). In the random validation, performances were overall higher, with R^2 up to 0.857 for the $SPAD + AGB$ model. Regarding UAV multispectral indices, the SIPI-based model showed the highest correlation with corn yield using LOTO validation (R^2 = 0.749), while integration with ground variables reduced errors, with minimal MAE and MAPE values in the $SIPI + SPAD + AGB$ model (MAE = 9.54 q ha^{-1}; MAPE = 11.99%). In random validation, the same combined model showed improved performance, with R^2 up to 0.824 and a significant reduction in errors (MAE = 6.36 q ha^{-1}; MAPE = 6.66%).

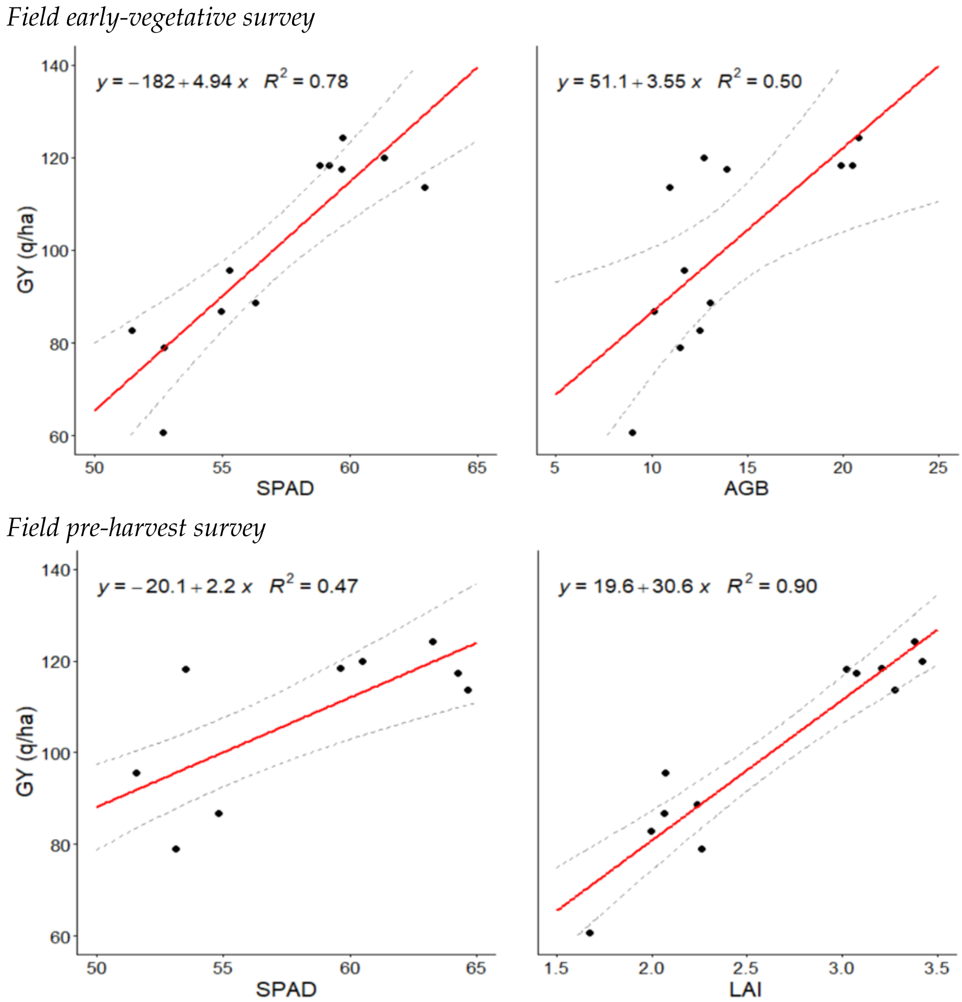

Figure 2. Relationships between grain yield (GY, q ha^{-1}) and ground-measured agronomic variables at different maize growth stages. The red solid lines represent the fitted linear regression models, while the grey dashed lines indicate the 95% confidence intervals.

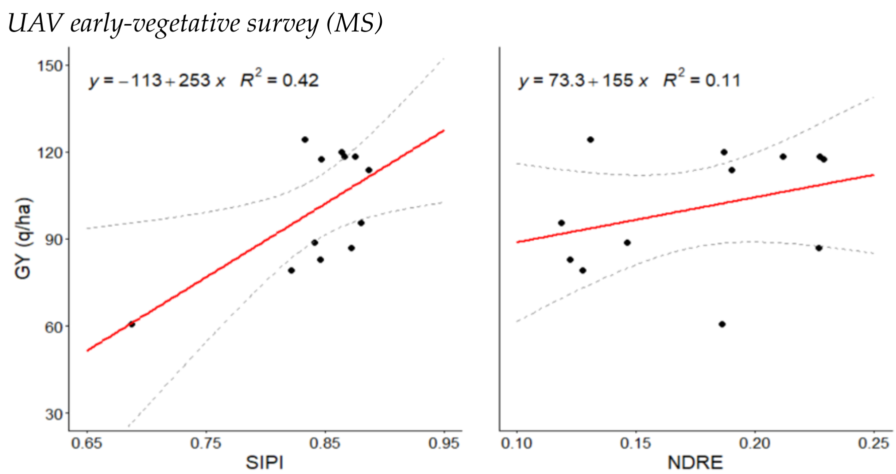

Figure 3. *Cont.*

UAV pre-harvest survey (MS)

UAV pre-harvest survey (HS)

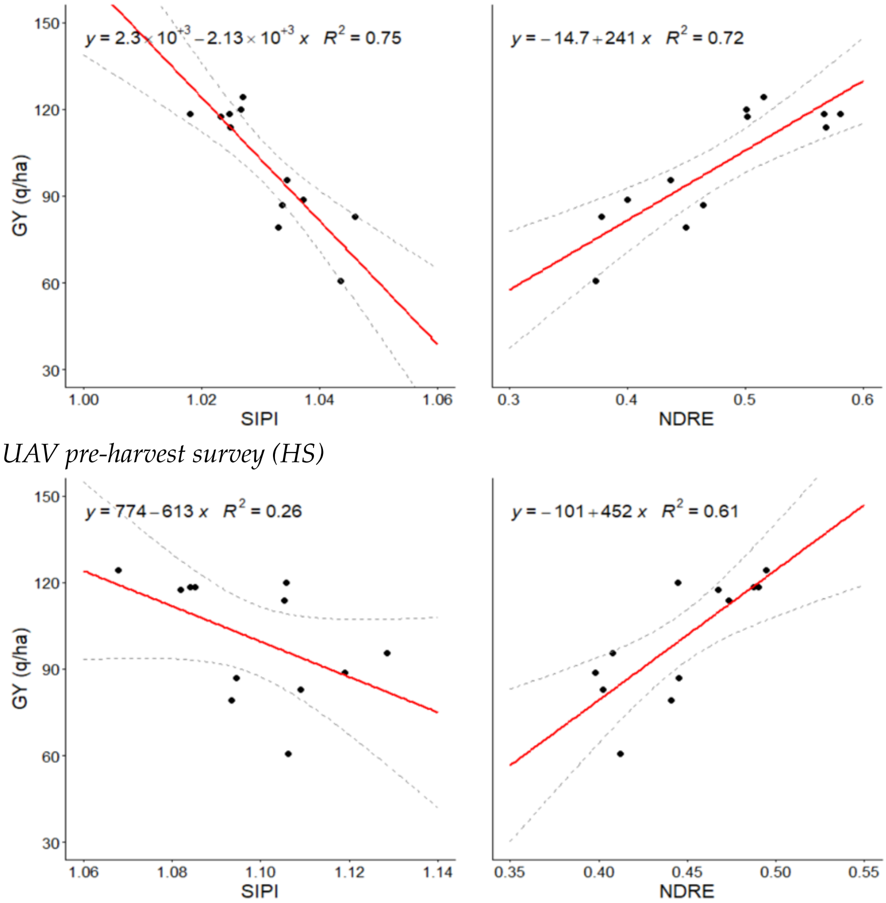

Figure 3. Relationships between grain yield (GY, q ha^{-1}) and UAV-derived Multispectral (MS) and Hyperspectral (HS) best indices at different maize growth stages. The red solid lines represent the fitted linear regression models, while the grey dashed lines indicate the 95% confidence intervals.

Table 4. Model performance for the early vegetative survey (July) using Leave-One-Treatment-Out (LOTO) and random validation.

Model	**LOTO**			**Random**		
	R^2	**MAE (q ha^{-1})**	**MAPE (%)**	**R^2**	**MAE (q ha^{-1})**	**MAPE (%)**
			Ground-based variables			
SPAD only	0.555	10.708	12.041	0.855	9.846	10.450
AGB only	0.455	14.999	16.674	0.591	13.592	14.271
SPAD + AGB	0.481	**8.236**	**9.544**	**0.857**	7.029	7.658
			Best UAV multispectral index			
SIPI only	**0.749**	26.095	30.154	0.571	16.839	17.204
SIPI + SPAD	0.639	15.766	19.633	0.821	10.868	11.861
SIPI + AGB	0.632	14.459	15.464	0.714	13.869	14.798
SIPI + SPAD + AGB	0.259	9.544	11.996	0.824	**6.364**	**6.656**

In the pre-harvest stage (September, Table 5), agronomic variables displayed generally superior performance compared to July. In the LOTO validation, the LAI-based model achieved R^2 = 0.699, while the combined $SPAD + LAI$ model showed a reduction in errors (MAE = 8.78 q ha^{-1}; MAPE = 10.25%). The best performance was observed for the LAI-only model using the random validation (R^2 = 0.870; MAE = 5.97 q ha^{-1}; MAPE = 6.59%). For UAV multispectral indices, the $SIPI + SPAD + LAI$ model achieved the best overall performance in LOTO validation (R^2 = 0.881), while in random validation the combined models generally presented high R^2 values (up to 0.863) and low errors (MAE = 6.57 q ha^{-1}; MAPE = 7.24%).

Table 5. Model performance for the pre-harvest survey (September) using Leave-One-Treatment-Out (LOTO) and random validation.

Model	**LOTO**			**Random**		
	R^2	**MAE (q ha^{-1})**	**MAPE (%)**	**R^2**	**MAE (q ha^{-1})**	**MAPE (%)**
			Ground-based variables			
SPAD only	0.493	12.008	12.957	0.736	11.197	11.752
LAI only	0.699	**8.125**	**9.570**	**0.870**	**5.974**	**6.592**
SPAD + LAI	0.493	8.776	10.253	0.853	6.839	7.455
			Best UAV multispectral index			
SIPI only	0.438	10.859	12.509	0.794	9.808	10.612
SIPI + SPAD	0.455	12.767	18.838	0.758	11.577	12.106
SIPI + LAI	0.659	8.168	9.595	0.858	6.569	7.236
SIPI + SPAD + LAI	**0.881**	9.424	10.878	0.863	7.517	8.145

Finally, the comparison between multispectral (MS) and hyperspectral (HS) data highlighted comparable performance between the two sensors (Table 6). In both datasets, the $NDRE + LAI$ model showed the best performance (R^2 = 0.870 for MS, and R^2 = 0.864 for HS), with comparable error values (MAE = 6.46 q ha^{-1} and MAE = 6.48 q ha^{-1}, respectively). The combined models with SPAD and LAI maintained similar levels of accuracy across the two data types but performed worse than the addition of LAI alone.

Table 6. Model performance using random validation of the NDRE index calculated from UAV multispectral and hyperspectral data acquired during the pre-harvest survey (September).

UAV Sensor	**Multispectral**			**Hyperspectral**		
	R^2	**MAE (q ha^{-1})**	**MAPE (%)**	**R^2**	**MAE (q ha^{-1})**	**MAPE (%)**
NDRE only	0.747	11.404	11.973	0.703	10.919	12.519
NDRE + SPAD	0.764	12.248	12.638	0.681	13.351	13.900
NDRE + LAI	**0.870**	**6.456**	**7.054**	**0.864**	**6.858**	**7.425**
NDRE + SPAD + LAI	0.860	8.120	8.601	0.822	9.078	9.370

3.3. Relative Importance of Predictors

To improve the interpretability of the best-performing models, standardized regression coefficients (β) were calculated for the predictors included in the most accurate models identified for each dataset (Table 7).

In the early-vegetative model ($SIPI + SPAD + AGB$), SPAD showed the largest standardized coefficient ($\beta = 0.61$, $p < 0.001$), followed by AGB ($\beta = 0.38$, $p = 0.002$) and SIPI ($\beta = 0.23$, $p = 0.032$).

In contrast, the pre-harvest models revealed a clear predominance of LAI. For both multispectral and hyperspectral datasets, LAI exhibited the highest standardized coefficients ($\beta = 0.83$ and 0.91, respectively) and was the only statistically significant predictor ($p < 0.05$). Conversely, NDRE and SPAD showed very small standardized coefficients and non-significant p-values, indicating a negligible independent contribution once canopy structural information was included in the model.

Table 7. Standardized regression coefficients (β) and associated p-values for the best-performing models identified in each dataset.

Dataset	Predictor	β	p-Value
July MS	SIPI	0.229	0.0315
	SPAD	0.610	<0.001
	AGB	0.376	<0.001
September MS	NDRE	0.124	0.5931
	SPAD	0.012	0.9603
	LAI	0.830	0.0167
September HS	NDRE	0.107	0.8746
	SPAD	0.040	0.8762
	LAI	0.907	<0.01

These results suggest that the relative importance of predictors changes substantially across phenological stages, with physiological indicators playing a stronger role during early growth, whereas canopy structure becomes the dominant determinant of yield prediction close to harvest.

4. Discussion

4.1. Relationship Between Grain Yield and Vegetation Indices

Grain yield (GY) prediction is a key element in supporting crop agronomic management, enabling optimized fertilization, irrigation, and harvesting strategies. To this end, numerous studies have used agronomic variables such as above-ground biomass (AGB), leaf chlorophyll content (SPAD), and leaf area index (LAI) as good predictors of productivity. AGB is directly linked to assimilate accumulation and consequently to the crop's productive potential; SPAD provides an indirect measure of nitrogen status and photosynthetic efficiency, while LAI describes the canopy's ability to intercept photosynthetically active radiation, regulating photosynthesis and growth processes [15,16]. In this study, significant relationships were observed between these parameters and corn grain yield (GY), with Pearson correlation coefficients ranging from $r \approx 0.5$ to 0.9, depending on the phenological stage (Figure 2), in line with previously reported correlation ranges for maize crops [16–18]. However, directly measuring these parameters requires intensive sampling campaigns, which are often destructive and difficult to scale across large areas. In this context, proximal and remote sensing represent an effective alternative, allowing crop productivity to be indirectly estimated through the analysis of plant spectral response.

Vegetation indices (VIs), in particular, represent one of the most commonly used tools for crop yield forecasting. VIs are mathematical combinations of spectral bands, generally in the visible and near-infrared, designed to amplify the vegetation signal and reduce the influence of confounding factors such as soil, lighting, and acquisition geometry [6]. Thanks to their computational simplicity and the possibility of being rapidly derived from data acquired by proximal sensors, UAVs or satellites, VIs provide indirect but reliable information on crop status necessary for precision agriculture applications [7]. Specifically, each index is linked to certain biophysical properties of vegetation; indices based on red-edge and green bands are more closely related to chlorophyll content and nutritional status, while those based on NIR and red bands are more sensitive to canopy structure and biomass. This functional diversity makes VIs particularly versatile tools for yield estimation but also means that their effectiveness strongly depends on phenological phases and growing conditions. For this reason, several recent studies have highlighted how selecting the most appropriate index, or combining multiple indices, is crucial to improve the accuracy of models for yield prediction [12,19,20].

4.2. Mechanistic Interpretation of Stage-Dependent Prediction Performance

The stronger correlations observed at the pre-harvest stage (R4) compared to the vegetative stage (V7) can be explained by the progressive accumulation of biomass and the establishment of canopy structure. At early stages, final yield is still poorly determined, and spectral signals are dominated by soil background and uneven cover. At later stages, the canopy is closed, and red-edge indices become sensitive to total chlorophyll content and nitrogen status, which directly relate to grain filling. This mechanistic interpretation is consistent with findings in maize [12,19] and other cereals [20].

In the present study, a correlation analysis was performed to identify VIs most strongly related to maize GY, and to assess potential redundancy between remotely sensed indices and ground-based measurements. The indices reported in Table 3 were then selected as predictors for subsequent modeling, not only based on the strength of their correlation with yield but also on their physiological significance and their ability to represent complementary biophysical processes of the crop. The results clearly highlighted that both the strength and direction of the relationships between VIs and GY were strongly dependent on corn phenological stage. During the vegetative phase (July UAV-MS campaign), relationships were weak and scattered, with only SIPI showing a significant correlation with GY ($r = 0.647$, $p = 0.02$) (Table 3 and Figure 3). This behavior can be attributed to the fact that, at early growth stages, final yield is still poorly determined, and the analyzed spectral response is mainly influenced by factors not directly related to productivity, such as soil cover and canopy structural variability [21]. In this context, more traditional VIs (e.g., NDVI) may be inefficient, whereas indices such as SIPI, sensitive to the carotenoid/chlorophyll ratio, are particularly effective in capturing early physiological variations.

Conversely, during the pre-harvest phase (September UAV-MS campaign), correlations became stronger and more clearly defined. Indices such as NDRE ($r = 0.849$, $p < 0.01$) and NDVI ($r = 0.843$, $p < 0.01$) showed a strong positive correlation with maize grain yield (Table 3 and Figure 3). This is consistent with previous studies [19,22], which demonstrated that in the later stages of the crop cycle, yield is strongly linked to accumulated biomass and canopy photosynthetic activity. Similarly, Fan et al. [12] reported that red-edge indices became strongly correlated with maize yield only from the mid-vegetative stage onward, while Ramos et al. [19] found that random forest models using NDRE achieved high accuracy at later stages. Our results extend these findings by quantifying the magnitude of improvement when ground data (LAI) are added to spectral indices. In particular, red-edge and green-based indices, such as GNDVI and NDRE, have often proven to outperform NDVI, as they are less affected by saturation and remain sensitive to chlorophyll and nitrogen content under high canopy density conditions [23,24]. Additionally, strong negative correlations were observed for SIPI and TVI ($r = -0.865$ and $r = -0.838$; $p < 0.01$), suggesting that these VIs are primarily driven by senescence processes and pigment degradation, resulting in an inverse relationship with yield at later crop stages. For this reason, such indicators were retained in the selection phase, as they can provide complementary information compared to similar VIs with opposite trends. Similarly, mCARI, despite showing weak correlations, was included due to its specific sensitivity to chlorophyll content and its variation [25], may allow the model to capture physiological components not fully represented by the other predictors.

Hyperspectral data acquired during the UAV pre-harvest survey further confirmed these findings, with red-edge indices such as VOG1 and GM2 showing the strongest positive correlations with maize yield ($r = 0.780$ and $r = 0.655$; $p < 0.01$), while TVI and SIPI again displayed negative relationships (Table 3). However, the comparable performance between hyperspectral and multispectral indices suggests that the presence of red-edge bands in multispectral sensors may already provide sufficient information for reliable yield estimation. Accordingly, several studies have highlighted how multispectral data can achieve accuracies very similar to hyperspectral data in estimating biophysical variables, such as chlorophyll (R^2 = 0.80 vs. 0.81, respectively), without substantial benefits deriving from the additional high spectral resolution [26,27].

4.3. Early Prediction of Grain Yield

The results of the early-vegetative (V7) estimation of maize grain yield (GY) highlighted how, in the initial growth stages, the predictive power of individual parameters is still limited (Table 4).

Regarding the ground-based variables, SPAD alone explained a good share of the yield variability (R^2 = 0.56–0.86) but with considerable errors (MAE = 10–12 q ha^{-1}, MAPE = 10–12%), while the use of AGB alone as a predictor led to even poorer performances (R^2 = 0.46–0.59, MAE = 14–16 q ha^{-1}). These findings indicate that, in early prediction, neither the state of leaf pigments nor the accumulation of biomass, considered individually, are sufficient to represent the variability that will subsequently characterize the final crop yield. A different picture emerged when considering the combination of the two agronomic variables. The combined model $SPAD + AGB$, in fact, while maintaining R^2 values comparable to the models based on single predictors (≈0.48–0.86), allowed a significant reduction in estimation errors, with MAE values around 7.5 q ha^{-1}. This behavior appears consistent from a biophysical point of view, suggesting that the maize yield potential in the early stages is the result of the interaction between canopy development and photosynthetic capacity, rather than the predominance of a single component. These results are in line with [20], which highlighted how the combination of variables related to the structure and physiological state of the crop (LAI and SPAD) improves the predictive ability of rice yield compared to the use of a single parameter (R^2 = 0.622 and 0.569, respectively). Similarly, the integration of the same agronomic variables allows for more accurate maize yield estimates than models based on single indices as demonstrated in [25].

The same trend was observed in predictions with multispectral indices calculated by UAVs (Table 4). In particular, the combination of $SIPI + SPAD + AGB$ significantly improved maize yield estimates, ranging from MAE values around 18–26 q ha^{-1} in the SIPI-only model to values between 6 and 9 q ha^{-1} for the combined model. This result highlights how spectral information alone is not sufficient to provide an accurate early prediction of GY but can significantly contribute when integrated with agronomic variables measured at ground level. Specifically, while SIPI is sensitive to variations in pigment content and hence to early physiological dynamics, the inclusion of SPAD and AGB allows for the integration of complementary information relating to the crop nutritional status and development. This interpretation is further supported by the standardized regression coefficients, which identified SPAD as the dominant predictor in the best-performing early-season model ($\beta = 0.61$), followed by AGB ($\beta = 0.38$) and SIPI ($\beta = 0.23$). These findings suggest that crop nitrogen status exerted a stronger influence on yield variability than either biomass accumulation or spectral information at this stage.

Overall, these findings suggest that, in the vegetative stages of maize growth, the most effective approach for yield estimation is the integration of proximal and remotely sensed data, rather than the use of a single indicator.

4.4. Late Prediction of Grain Yield

Pre-harvest yield forecasts showed overall superior performance compared to those obtained in the early vegetative stage. However, unlike what was observed for early estimates, the integration of variables did not produce equally significant improvements (Table 5).

https://doi.org/10.3390/rs18152460

In particular, among the ground-based variables, LAI used as a single predictor showed the best performance (R^2 = 0.70–0.90; MAE = 6–8 q ha^{-1}; MAPE = 7–10%), outperforming both the SPAD-only (R^2 = 0.50–0.75) model and the combined $SPAD + LAI$ model (R^2 = 0.50–0.85). This result suggests that, in the later stages of the crop cycle, yield is more closely linked to canopy structure and radiation interception capacity, rather than chlorophyll content, the role of which is relatively less significant once production potential is already established. Notably, the LAI-only model achieved performance comparable to the fused models ($R^2 = 0.87$, MAE $= 5.97$ q ha^{-1}), indicating that in some phenological phases a single well-chosen structural variable can be as effective as more complex combinations. This finding cautions against assuming that adding more features always improves prediction.

A similar pattern emerges from the analysis of UAV-derived multispectral indices. First, the performance of models based on single VIs is generally better at the late stage than at the early phenological stage. Moreover, in addition to SIPI, other indices, such as NDVI and NDRE, also showed good predictive capacity for grain yield in the pre-harvest survey (Table 6), confirming a greater stability of the relationship between spectral response and final yield at this phenological stage. In line with our findings, several recent studies have highlighted how the predictive power of vegetation indices and agronomic variables differs significantly as a function of crop phenological stage and experimental conditions. For example, [13] showed that indices based on NIR and Red-Edge bands are among the best predictors of maize yield, with peak performance for mid-season acquisitions. Similarly, [28] highlighted how the relationship between VIs and yield varies as a function of agronomic treatments and growth stage, with some indices (including Red-Edge indices) showing good predictive power across different conditions, while the inclusion of uninformative variables can significantly reduce model performance.

Regardless of the acquisition date, integrating proximal and remotely sensed data remains an effective strategy. Combined models, such as $SIPI + LAI$, produced more accurate estimates (R^2 = 0.86–0.88; MAE = 7–9 q ha^{-1}; MAPE = 8–10%), further reducing errors compared to using individual predictors. The importance of integrating structural and physiological variables has previously been highlighted by [20], which reported improved yield estimates through the combination of parameters such as LAI and SPAD compared to their individual use, also observing a reduction in the contribution of chlorophyll content for late prediction. Nevertheless, data fusion does not always improve performance and may introduce redundancy or noise when predictors are highly correlated. For instance, in the pre-harvest stage, adding SPAD to the LAI-only model did not reduce errors (Table 5), indicating that LAI already captured most of the yield-relevant information. This interpretation is strongly supported by the standardized regression coefficients of the best-performing models. In both the multispectral and hyperspectral datasets, LAI showed the highest contribution to yield estimation ($\beta = 0.83$ and 0.91, respectively) and was the only statistically significant predictor. Conversely, the contribution of NDRE and SPAD became negligible once LAI was included in the model.

Overall, these findings suggest that, at the pre-harvest stage, canopy structural traits already integrate much of the information related to crop vigor and productivity, reducing the additional predictive value of spectral and chlorophyll-based indicators.

4.5. Validation Strategy Dependence

Comparing the two validation strategies highlighted systematic differences in model performance. Overall, random validation produced higher R^2 values and lower errors than Leave-One-Treatment-Out (LOTO) validation. This behavior was observed in all model configurations and for both phenological phases analyzed (Tables 4 and 5).

These differences can be attributed to the dissimilar structure of the datasets used in the two approaches. In the case of random validation, samples belonging to the same treatment may be present in both the training and test sets, introducing a dependency between the data that tends to favor apparently higher performance. In contrast, LOTO validation represents a more restrictive scenario, in which the model is tested on conditions completely unobserved during training, providing a more realistic estimate of generalization ability. This aspect is particularly relevant in the present study, characterized by a limited number of treatments but with marked differences between them. Under these conditions, random validation can lead to an overestimation of predictive accuracy, while LOTO validation allows for a more robust assessment of model performance across the treatments included in the experimental design.

Overall, these results suggest that the choice of validation strategy is a crucial element in the evaluation of yield prediction models. In particular, approaches that take into account the experimental structure, such as LOTO validation, are more appropriate for estimating the actual predictive capability in operational scenarios. Similarly, [13] highlighted how the performance of yield estimation models is strongly influenced by the validation strategy adopted, with independent approaches producing more realistic but less optimistic estimates. This behavior is also consistent with what [8] reported in a recent review paper, where the authors emphasizes that the generalization of yield forecasting models represents a major challenge in studies based on UAV data, due to the strong dependence on experimental conditions and the limited availability of representative datasets.

4.6. Multispectral vs. Hyperspectral Data

The comparison between multispectral (MS) and hyperspectral (HS) data revealed no substantial differences in the performance of yield estimation models. Overall, the two data types showed comparable predictive power in terms of coefficient of determination and estimation errors (Table 6).

In particular, indices based on the red-edge region, such as NDRE, provided similar results regardless of the sensor type used. This suggests that the presence of targeted spectral bands, such as the red-edge and near-infrared, in multispectral sensors is already sufficient to capture key biophysical information related to yield, including chlorophyll content and crop nutritional status. Consistently, [22] observed comparable performance between multispectral and hyperspectral data in wheat yield estimation, with similar R^2 values for both sensor types. The authors also highlighted how red-edge-based indices represent the best predictors, thanks to their greater sensitivity to chlorophyll and nitrogen content and lower saturation at high leaf densities.

Overall, the MS vs. HS comparison suggests that the use of multispectral data equipped with a red-edge band represents a practical compromise between accuracy and operational complexity under the conditions tested. However, the inclusion of ground-based measurements (SPAD, LAI, AGB) introduces operational constraints (time, labour, destructiveness) that limit real-time applicability. Therefore, the proposed approach is best suited for research or decision-support contexts where ground data are already collected, rather than for fully automated high-throughput prediction. We therefore reframe the contribution as a mechanistic dissection of yield formation rather than a purely rapid prediction tool. However, the advantage of hyperspectral imaging is greatest in fields or conditions with poor uniformity or in the presence of stress, where multispectral sensors cannot discriminate well [29].

4.7. Study Limitations and Future Research

Several limitations should be acknowledged. First, the dataset is small (n = 12 plots) and derived from a single growing season and a single location, which limits the statistical power and external validity. Second, only one maize hybrid was used, and the fertilization treatments, while realistic, do not capture the full range of agronomic practices. Third, the comparison between multispectral and hyperspectral sensors was not fully balanced (different acquisition dates, spatial resolutions, and pre-processing workflows). Fourth, the hyperspectral data were used only through a limited set of vegetation indices; full-spectrum methods (e.g., partial least squares, machine learning on all bands) were not explored. These aspects represent limitations of the present dataset and identify priorities for future research rather than conclusions supported by the current study. Future research should address these gaps by extending the experiment to multiple years, sites, and genotypes. Advanced methods such as domain adaptation, transfer learning, or spatial cross-validation represent promising directions for future research but were beyond the scope of this study. Their potential contribution to broader model generalization remains to be established through dedicated multi-site and multi-year studies. Future work should explore these techniques to determine whether they improve model robustness following appropriate validation across multiple sites and growing seasons. Specifically, future studies should further exploit the full spectral richness of hyperspectral data through dimensionality reduction (e.g., principal component analysis, partial least squares regression), band selection, or full-spectrum modeling approaches, which may help reduce redundancy while preserving the most informative spectral features for yield prediction.

5. Conclusions

The present study investigated the potential of integrating ground-based measurements and UAV-derived spectral data for predicting maize grain yield at different phenological stages and under different nitrogen treatments.

The results showed a strong dependence of model performance on the crop's growth stage. Yield prediction at the early vegetative stage performed poorly and benefited significantly from the integration of complementary variables, indicating that no single predictor was sufficient to capture the variability of final yield. Predictive performance improved overall in the pre-harvest stage, especially for vegetation indices, with LAI proving to be the best ground-based variable to integrate with UAV data.

Comparison of multispectral and hyperspectral data showed comparable performance, suggesting that multispectral sensors equipped with red-edge bands are sufficient to capture key spectral information related to yield, providing a cost-effective and operationally efficient alternative to hyperspectral systems in similar contexts.

Furthermore, analysis of different validation strategies demonstrated that random sampling tends to overestimate model performance, while the Leave-One-Treatment-Out approach provides a more realistic assessment of model behavior across the treatments included in the experimental design.

Overall, our findings confirm that yield estimation is the result of a balance between different biophysical factors, whose relative importance varies over time, and that the effectiveness of predictive models strongly depends on both the selection of variables and the phenological phase in which observations are acquired.

Nevertheless, these findings should be interpreted as a proof-of-concept, as they are based on a single-site, single-season experiment with a limited sample size (n = 12). Although the proposed workflow showed promising results under the investigated conditions, further validation across multiple sites, growing seasons, and larger datasets is required before its broader applicability can be established. Future research should also investigate the use of full-spectrum hyperspectral information and dimensionality-reduction approaches, together with validation across multiple sites and growing seasons, to further assess the robustness and broader applicability of the proposed workflow.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs18152460/s1. Table S1. Vegetation Indices calculated from UAV-acquired multispectral data. Table S2. Vegetation Indices calculated from UAV-acquired hyperspectral data. in addition to those reported in Table S1. Table S3. Pearson correlation coefficients between vegetation indices (VIs) and grain yield (GY) for multispectral data acquired by UAV during the early vegetative (V7) stage of maize crop. The underlined VIs represent the selected predictors for subsequent maize yield prediction models considering both correlation strength and physiological relevance. Table S4. Pearson correlation coefficients between vegetation indices (VIs) and grain yield (GY) for multispectral data acquired by UAV during the pre-harvest (R4) stage of maize crop. The underlined VIs represent the selected predictors for subsequent maize yield prediction models considering both correlation strength and physiological relevance. Table S5. Pearson correlation coefficients between vegetation indices (VIs) and grain yield (GY) for hyperspectral data acquired by UAV during the pre-harvest (R4) stage of maize crop. The underlined VIs represent the selected predictors for subsequent maize yield prediction models considering both correlation strength and physiological relevance. Table S6. Model performance for the ground-based variables of the early vegetative (V7) and pre-harvest (R4) surveys using Leave-One-Treatment-Out (LOTO) and Random validation. Table S7. Model performance for the selected UAV-multispectral indices of the early vegetative (V7) survey using Leave-One-Treatment-Out (LOTO) and Random validation. Table S8. Model performance for the selected UAV-multispectral indices of the pre-harvest (R4) survey using Leave-One-Treatment-Out (LOTO) and Random validation. Table S9. Model performance for the selected UAV-hyperspectral indices of the pre-harvest (R4) survey using Leave-One-Treatment-Out (LOTO) and Random validation. Figure S1. Relationships between maize grain yield (GY. q/ha) and the selected UAV-multispectral indices of the early vegetative (V7) survey. Figure S2. Relationships between maize grain yield (GY. q/ha) and the selected UAV-multispectral indices of the pre-harvest (R4) survey. Figure S3. Relationships between maize grain yield (GY. q/ha) and the selected UAV-hyperspectral indices of the pre-harvest (R4) survey.

Author Contributions: Conceptualization, S.P., F.T., C.S. and M.G.; methodology, S.P., C.S. and M.G.; software, M.G.; validation, S.P., C.S. and M.G.; formal analysis, C.S. and M.G.; investigation, S.P., M.G. and C.S.; resources, V.C., D.S. and V.D.M.; data curation, F.T., C.S. and M.G.; writing—original draft preparation, C.S. and M.G.; writing—review and editing, S.P. and M.D.M.; supervision, S.P. and M.D.M.; project administration, S.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Italian Ministry of Universities and Research (MUR) under the National Aerospace Research Program (PRORA), project "Bioscienza (Terra-bioindicatori)", DM 662/2020, CUP I44G20000040001. This research was partially supported by the Project 'Green Farm', funded by the Italian Ministry of Economic Development (MiSE)-PON 'Imprese e Competitività' 2014-2020 FESR, Project No. F/200117/01/X45, CUP B66G20000660005.

Data Availability Statement: Data regarding biomass and yield are property of Green Farm Project (CUP B66G20000660005) and are available from Vincenza Cozzolino upon reasonable request; Data regarding non-destructive field measurements and UAV acquisitions are part of the "Bioscienza (Terra-bioindicatori)" project and are available from Massimiliano Gargiulo upon reasonable request.

Acknowledgments: The authors thank the Experimental Farm at Castel Volturno for facilities and the staff of Project GreenFarm for their essential contribution to the field trials. The authors used AI tools for minor language editing and text refinement. All scientific content and interpretations are the sole responsibility of the authors.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

AGB	above-ground biomass
DAS	days after sowing
GY	grain yield
HS	hyperspectral
LAI	leaf area index
LOTO	leave-one-treatment-out
MAE	mean absolute error
MAPE	mean absolute percentage error
MS	multispectral
NDVI	normalized difference vegetation index
NDRE	normalized difference red-edge
OLS	ordinary least squares
R^2	coefficient of determination
SPAD	soil plant analysis development
UAV	unmanned aerial vehicle
VI	vegetation index

References

1. Ray, D.K.; Mueller, N.D.; West, P.C.; Foley, J.A. Yield trends are insufficient to double global crop production by 2050. *PLoS ONE* **2013**, *8*, e66428. [CrossRef] [PubMed]
2. Gamage, A.; Gangahagedara, R.; Gamage, J.; Jayasinghe, N.; Kodikara, N.; Suraweera, P.; Merah, O. Role of organic farming for achieving sustainability in agriculture. *Farming Syst.* **2023**, *1*, 100005. [CrossRef]
3. Panday, D.; Bhusal, N.; Das, S.; Ghalehgolabbehbahani, A. Rooted in nature: The rise, challenges, and potential of organic farming and fertilizers in agroecosystems. *Sustainability* **2024**, *16*, 1530. [CrossRef]
4. Zhang, Q. *Precision Agriculture Technology for Crop Farming*; Taylor & Francis: Abingdon, UK, 2016.
5. Shafi, U.; Mumtaz, R.; García-Nieto, J.; Hassan, S.A.; Zaidi, S.A.R.; Iqbal, N. Precision agriculture techniques and practices: From considerations to applications. *Sensors* **2019**, *19*, 3796. [CrossRef] [PubMed]
6. Sishodia, R.P.; Ray, R.L.; Singh, S.K. Applications of remote sensing in precision agriculture: A review. *Remote Sens.* **2020**, *12*, 3136. [CrossRef]
7. Radočaj, D.; Šiljeg, A.; Marinović, R.; Jurišić, M. State of major vegetation indices in precision agriculture studies indexed in web of science: A review. *Agriculture* **2023**, *13*, 707. [CrossRef]
8. Yuan, J.; Zhang, Y.; Zheng, Z.; Yao, W.; Wang, W.; Guo, L. Grain Crop Yield Prediction Using Machine Learning Based on UAV Remote Sensing: A Systematic Literature Review. *Drones* **2024**, *8*, 559. [CrossRef]
9. Baio, F.H.R.; Santana, D.C.; Teodoro, L.P.R.; Oliveira, I.C.d.; Gava, R.; de Oliveira, J.L.G.; Silva Junior, C.A.d.; Teodoro, P.E.; Shiratsuchi, L.S. Maize yield prediction with machine learning, spectral variables and irrigation management. *Remote Sens.* **2022**, *15*, 79. [CrossRef]
10. Chatterjee, S.; Adak, A.; Wilde, S.; Nakasagga, S.; Murray, S.C. Cumulative temporal vegetation indices from unoccupied aerial systems allow maize (*Zea mays* L.) hybrid yield to be estimated across environments with fewer flights. *PLoS ONE* **2023**, *18*, e0277804. [CrossRef] [PubMed]
11. Liu, S.; Hu, Z.; Han, J.; Li, Y.; Zhou, T. Predicting grain yield and protein content of winter wheat at different growth stages by hyperspectral data integrated with growth monitor index. *Comput. Electron. Agric.* **2022**, *200*, 107235. [CrossRef]
12. Fan, J.; Zhou, J.; Wang, B.; de Leon, N.; Kaeppler, S.M.; Lima, D.C.; Zhang, Z. Estimation of maize yield and flowering time using multi-temporal UAV-based hyperspectral data. *Remote Sens.* **2022**, *14*, 3052. [CrossRef]
13. Killeen, P.; Kiringa, I.; Yeap, T.; Branco, P. Corn grain yield prediction using UAV-based high spatiotemporal resolution imagery, machine learning, and spatial cross-validation. *Remote Sens.* **2024**, *16*, 683. [CrossRef]
14. Gargiulo, M.; Savarese, C.; De Mizio, M.; Tufano, F.; Parrilli, S. Rapid Monitoring of Maize Plant Status Using UAV-Based Multispectral and Hyperspectral Data. In *Proceedings of the 2024 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor)*; IEEE: New York, NY, USA, 2024; pp. 7–12.

15. Gitelson, A.A.; Peng, Y.; Arkebauer, T.J.; Schepers, J. Relationships between gross primary production, green LAI, and canopy chlorophyll content in maize: Implications for remote sensing of primary production. *Remote Sens. Environ.* **2014**, *144*, 65–72. [CrossRef]
16. Berdjour, A.; Dugje, I.; Rahman, N.A.; Odoom, D.A.; Kamara, A.; Ajala, S. Direct estimation of maize leaf area index as influenced by organic and inorganic fertilizer rates in Guinea Savanna. *J. Agric. Sci.* **2020**, *12*, 66. [CrossRef]
17. Chen, K.; Camberato, J.J.; Vyn, T.J. Maize grain yield and kernel component relationships to morphophysiological traits in commercial hybrids separated by four decades. *Crop Sci.* **2017**, *57*, 1641–1657. [CrossRef]
18. Zagyi, P.; Tamás, A.; Rácz, D.; Fejér, P.; Radócz, L.; Horváth, É. Correlation analysis of relative chlorophyll content and yield of maize hybrids of different genotypes. *Acta Agrar. Debreceniensis* **2022**, *1*, 211–214. [CrossRef]
19. Ramos, A.P.M.; Osco, L.P.; Furuya, D.E.G.; Gonçalves, W.N.; Santana, D.C.; Teodoro, L.P.R.; da Silva Junior, C.A.; Capristo-Silva, G.F.; Li, J.; Baio, F.H.R.; et al. A random forest ranking approach to predict yield in maize with uav-based vegetation spectral indices. *Comput. Electron. Agric.* **2020**, *178*, 105791. [CrossRef]
20. 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]
21. Xue, J.; Su, B. Significant remote sensing vegetation indices: A review of developments and applications. *J. Sens.* **2017**, *2017*, 1353691. [CrossRef]
22. Liu, Y.; Sun, L.; Liu, B.; Wu, Y.; Ma, J.; Zhang, W.; Wang, B.; Chen, Z. Estimation of Winter Wheat Yield using multiple temporal vegetation indices derived from UAV-Based multispectral and hyperspectral imagery. *Remote Sens.* **2023**, *15*, 4800. [CrossRef]
23. Macedo, F.L.; Nóbrega, H.; de Freitas, J.G.; Ragonezi, C.; Pinto, L.; Rosa, J.; Pinheiro de Carvalho, M.A. Estimation of productivity and above-ground biomass for corn (*Zea mays*) via vegetation indices in Madeira Island. *Agriculture* **2023**, *13*, 1115. [CrossRef]
24. Sapkota, S.; Paudyal, D.R. Growth monitoring and yield estimation of maize plant using unmanned aerial vehicle (UAV) in a hilly region. *Sensors* **2023**, *23*, 5432. [CrossRef] [PubMed]
25. 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]
26. Tong, A.; He, Y. Estimating and mapping chlorophyll content for a heterogeneous grassland: Comparing prediction power of a suite of vegetation indices across scales between years. *ISPRS J. Photogramm. Remote Sens.* **2017**, *126*, 146–167. [CrossRef]
27. Lu, B.; He, Y.; Dao, P.D. Comparing the performance of multispectral and hyperspectral images for estimating vegetation properties. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 1784–1797. [CrossRef]
28. Kumar, C.; Mubvumba, P.; Huang, Y.; Dhillon, J.; Reddy, K. Multi-stage corn yield prediction using high-resolution UAV multispectral data and machine learning models. *Agronomy* **2023**, *13*, 1277. [CrossRef]
29. Yang, C.; Everitt, J.; Bradford, J.; Murden, D. Comparison of airborne multispectral and hyperspectral imagery for estimating grain sorghum yield. *Trans. ASABE* **2009**, *52*, 641–649. [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.