agronomy

MDPI

Article

Prediction of Winter Wheat Parameters with Planet SuperDove Imagery and Explainable Artificial Intelligence

Gabriele De Carolis [1,2], Vincenzo Giannico [2], Leonardo Costanza [2], Francesca Ardito [2], Anna Maria Stellacci [2], Afwa Thameur [3], Sergio Ruggieri [1], Sabina Tangaro [2], Marcello Mastrorilli [1], Nicola Sanitate [1] and Simone Pietro Garofalo [1,*]

[1] Council for Agricultural Research and Economics, Research Center for Agriculture and Environment, Via Celso Ulpiani, 5, 70125 Bari, Italy
[2] Department of Soil, Plant and Food Science, University of Bari Aldo Moro, Via Amendola 165/A, 70126 Bari, Italy
[3] Laboratory of Biodiversity, Molecules, Application (BMA), Higher Institute of Applied Biology Medenine, University of Gabes, 4119 Medenine, Tunisia
* Correspondence: simone.garofalo@crea.gov.it

Abstract: This study investigated the application of high-resolution satellite imagery from SuperDove satellites combined with machine learning algorithms to estimate the spatiotemporal variability of some winter wheat parameters, including the relative leaf chlorophyll content (RCC), relative water content (RWC), and aboveground dry matter (DM). The research was carried out within an experimental field in Southern Italy during the 2024 growing season. Different machine learning (ML) algorithms were trained and compared using spectral band data and calculated vegetation indices (VIs) as predictors. Model performance was assessed using R^2 and RMSE. The ML models tested were random forest (RF), support vector regressor (SVR), and extreme gradient boosting (XGB). RF outperformed the other ML algorithms in the prediction of RCC when using VIs as predictors (R^2 = 0.81) and in the prediction of the RWC and DM when using spectral bands data as predictors (R^2 = 0.71 and 0.87, respectively). Model explainability was assessed with the SHAP method. A SHAP analysis highlighted that GNDVI, Cl1, and NDRE were the most important VIs for predicting RCC, while yellow and red bands were the most important for DM prediction, and yellow and nir bands for RWC prediction. The best model found for each target was used to model its seasonal trend and produce a variability map. This approach highlights the potential of integrating ML and high-resolution satellite imagery for the remote monitoring of wheat, which can support sustainable farming practices.

Keywords: chlorophyll content; relative water content; dry matter; remote sensing; planet; machine learning; explainable artificial intelligence

check for updates

Academic Editor: Maofang Gao

Received: 13 December 2024
Revised: 3 January 2025
Accepted: 17 January 2025
Published: 19 January 2025

Citation: De Carolis, G.; Giannico, V.; Costanza, L.; Ardito, F.; Stellacci, A.M.; Thameur, A.; Ruggieri, S.; Tangaro, S.; Mastrorilli, M.; Sanitate, N.; et al. Prediction of Winter Wheat Parameters with Planet SuperDove Imagery and Explainable Artificial Intelligence. *Agronomy* **2025**, *15*, 241. https://doi.org/10.3390/agronomy15010241

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

1. Introduction

1.1. Background

Agriculture and food systems are facing increasingly complex challenges, such as the need to increase food production by 25–70% over current levels while maintaining or reducing the inputs used [1]. Generally, many farmers manage their fields without properly considering the spatial diversity, such as soil parameters, growing conditions, and/or other agronomic factors [2,3]. Thus, the result is often the improper use of agrochemicals, fossil fuels, and machinery, with a negative impact on the agroecosystem [4]. Sustainable intensification is emerging in the current debate on the future of agriculture. The aim is to

increase food production by improving efficiency, reducing waste, preserving resources, and minimizing environmental impacts [1,5]. Some important approaches include improving agronomic techniques and introducing new varieties with higher performance [6]. Studying the temporal and spatial distribution of the physiological characteristics of a crop can be useful in adopting the best agronomic techniques for optimizing yield and improving crop sustainability. Winter wheat is one of the most important food crops in the world, with a significant economic relevance [7], providing 20% of the calories in the human diet. It is a source of fiber, vitamins, minerals, and bioactive compounds (polyphenols and carotenoids) [8–11]. Europe is the second most important continent after Asia for wheat production. Italy is one of the first producers with an average cultivated area of 1.23 million hectares [12]. Over the last half-century, intensive wheat cultivation has increased yields and economic profits but has led to environmental concerns [13]. For instance, the excessive use of fertilizers could cause nutrient leaching and the emissions of greenhouse gases such as nitrogen dioxide and nitrous oxide [14]. However, climate change may lead to negative effects with potential losses in wheat productivity, particularly under heat and drought conditions [15–17].

1.2. Remote Sensing for Wheat Monitoring

To improve the sustainability of cereal production, several approaches have been studied, including remote sensing techniques for the monitoring of key agronomic variables [18]. Many research works have focused on remote monitoring of chlorophyll content, which is an indicator of crop health and nutritional status. Monitoring chlorophyll levels can allow early intervention to optimize nitrogen fertilization [19–21]. Yang et al. [22] demonstrated that drought significantly dropped both the chlorophyll content and leaf photosynthetic efficiency (Fv/Fm ratio) across various growth stages. Moreover, the total chlorophyll content was reported to be positively correlated with thousand kernel weight under drought conditions [23]. Through the use of remote sensing technologies, researchers, farmers, and decision-makers could efficiently obtain information on various crop characteristics [24,25]. In addition, the integration of data from multiple sources, such as satellite platforms, unmanned aerial vehicle (UAV) data, and ground-based sensors, can improve the accuracy of agronomic predictions by supporting the continuous monitoring on a regional and field scale and providing detailed information on crop growth dynamics. Remote sensing data and derived vegetation indices, such as the Normalized Difference Vegetation Index, are widely used to estimate plant characteristics [26]. Different satellite platforms have been used to monitor crop parameters, including Landsat and Sentinel-2, which offer moderate spatial resolutions of 30 and 10 m, respectively, and revisit periods ranging from 5 to 16 days. In contrast, SuperDove satellites, part of the PlanetScope constellation, offer a higher spatial resolution (3 m) and a daily revisit frequency, potentially leading to more detailed and timely monitoring of the growth and status of wheat [27]. However, it is important to consider that while SuperDove offers higher spatial resolution, Sentinel-2 provides a broader spectral range with 13 bands, including three red edge bands and shortwave infrared bands, which could be particularly useful for wheat monitoring [28]. Ogungbuyi et al. [29] estimated the vegetation biomass by combining the "SPECTRA-FOR" (Spectral Pasture Estimation using Combined Techniques of Random Forest Algorithm for Feature Optimization and Retrieval) algorithm and satellite data from both Sentinel-2 and SuperDove. Their approach highlighted the potential advantages of integrating high-resolution satellite imagery with advanced algorithms for more precise biomass estimation. The integration of spectral data from SuperDove's images improved the model's performance, increasing the R^2 value from 0.87 (Sentinel-2) to 0.92 (SuperDove) and reducing the MSE from 439 to 346 kg/ha.

Machine learning approaches, using algorithms such as Random Forest, Support Vector Machines, and Neural Networks, are useful for predicting biophysical parameters from remote sensing data, providing innovative solutions to agricultural challenges [30–38].

This work aims to remotely monitor different winter wheat features (the relative leaf chlorophyll content, relative water content, and aboveground dry matter) at the field scale using spectral information from high-resolution satellites (Planet) and comparing different machine learning algorithms for regressions, including random forest, support vector regressor, and extreme gradient boosting. These algorithms were selected because they are considered robust and effective in analyzing agricultural datasets, which are often characterized by complex patterns and/or non-normal distributions [39,40]. The machine learning algorithms used in this study do not assume a specific distribution of the data thus making them particularly suited for various types of datasets.

2. Materials and Methods

2.1. Study Area and Climate Conditions

The experiment was conducted in Rutigliano, south of Italy (40°59′ N; 17°01′ E; 125 m above sea level), within the experimental farm belonging to the Council for Agricultural Research and the Economic Analysis (CREA) (Figure 1A). The soil of the experimental farm was classified as clay loam (21% sand, 37% silt, and 42% clay). The climate around the area is warm and temperate with mild winters and hot summers, and it is classified as CSa according to Köppen and Geiger [41]. Most of the rainfall occurs during the autumn and winter seasons, and summers are usually dry. The average annual precipitation is about 622 mm, the average annual temperature is 16.5 °C, the coldest month is January (average temperature: 8.7 °C), and the hottest month is August (average temperature: 25.5 °C) [42]. The field-specific agrometeorological data for the year of the experiment were collected from the agrometeorological station installed within the experimental farm; the daily crop evapotranspiration was calculated following the FAO-56 methodology [43].

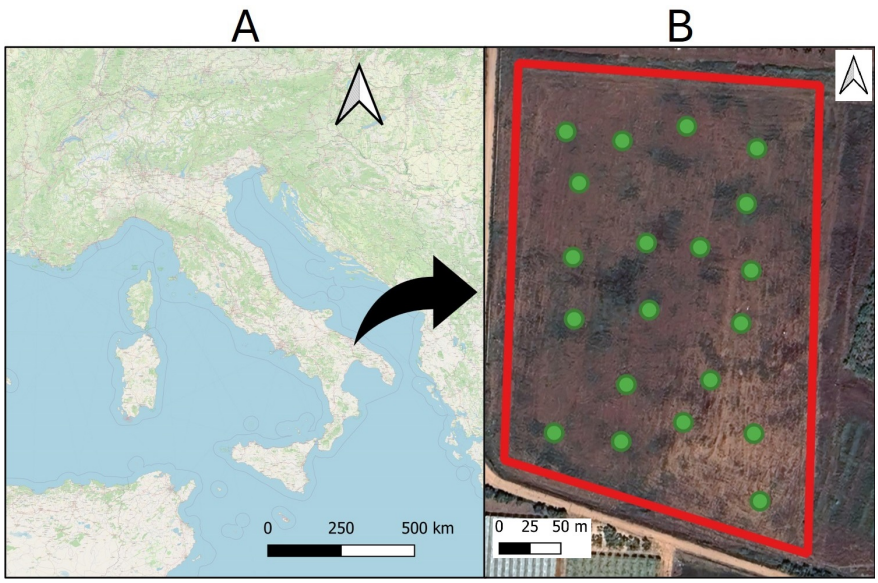


Figure 1. The location of the experimental farm, south of Italy (**A**); random points (green dots) where the relative leaf chlorophyll content, relative water content, and aboveground dry matter were determined within the field (red line) (**B**) (OpenStreetMap Contributors, 2024; Map data© 2015 Google).

2.2. Field Data

Wheat (*Triticum turgidum*, ssp. durum, Desf. , cv 'Antalis') was sown on 16 December 2023 on an area of 1.7 ha with a density of 200 seeds m^{-2}. Before this cultivation process, the field was left fallow and uniformly fertilized during pre-sowing with 150 kg/ha of ammonium nitrate and 100 kg/ha of triple superphosphate in accordance with the 2024 guidelines of the Integrated Production Discipline of Puglia. These guidelines ensure that fertilization practices meet sustainability criteria, optimizing nutrient use to support crop growth while minimizing environmental impacts.

The parameters acquired and considered as targets to predict in this study were the relative leaf chlorophyll content (RCC), relative water content (RWC), and aboveground dry matter (DM). The RCC, RWC, and DM were acquired on 20 random points within the field, and for each point, 3 replicates were determined and averaged (Figure 1B). All of these parameters were acquired under cloud-free conditions. The RCC was measured by using a hand-held chlorophyll meter (MC-100, Apogee Instruments, Inc. 721 W 1800 N Logan, Utah 84321, USA) from the tillering phase (end of February) to pre-harvest (May) at 7 sampling times. The instrument provides an estimation of the actual leaf chlorophyll concentration expressed in μmol m^{-2}. The RWC was determined from the stem elongation phase (March) to pre-harvest (May) on the youngest fully expanded leaf [44] at 5 sampling times. RWC was calculated as follows (1):

$$RWC(\%) = \frac{(FW - DW)}{(TW - DW)} \times 100 \quad (1)$$

where FW is the fresh weight of the leaf, TW is the turgid weight of the leaf (obtained after placing the leaf in distilled water for 16 h), and DW is the dry weight of the leaf (measured after 48 h at 65 °C in the oven). DM was determined from the stem elongation phase to pre-harvest at 6 sampling times using the following equation:

$$DM(\%) = \frac{DW}{FW} \times 100 \quad (2)$$

where FW is the fresh weight of the samples and DW is the dry weight of the samples after 48 h in an oven at 48 °C.

A summary of statistics of the ground truth data for each parameter is reported in Table 1. RCC ranged between 8.53 μmol m^{-2} and 553.93 μmol m^{-2} and had a mean value of 321.31 μmol m^{-2}. RWC ranged between 12.37% and 93.75%, and the average was 68.07%. DM ranged between 22.74% and 65.73%, and the average DM was 44.56%. Figure 2 reports the distribution frequency of RCC, RWC, and DM and their kurtosis and skewness values. All of the parameters considered showed a distribution different from the normal, with kurtosis values lower than 3 and skewness values different from 0.

Table 1. A summary of the statistics of the ground truth parameters acquired for this study (relative leaf chlorophyll content, RCC; relative water content, RWC; aboveground dry matter, DM). s.d. = standard deviation.

Parameter	Count	Mean	s.d.	Median	3rd Quartile	Min	Max
RCC	140	321.31	142.95	351.53	432.03	8.53	553.93
RWC	100	0.681	0.196	0.717	0.851	0.123	0.937
DM	120	0.445	0.119	0.442	0.555	0.227	0.657

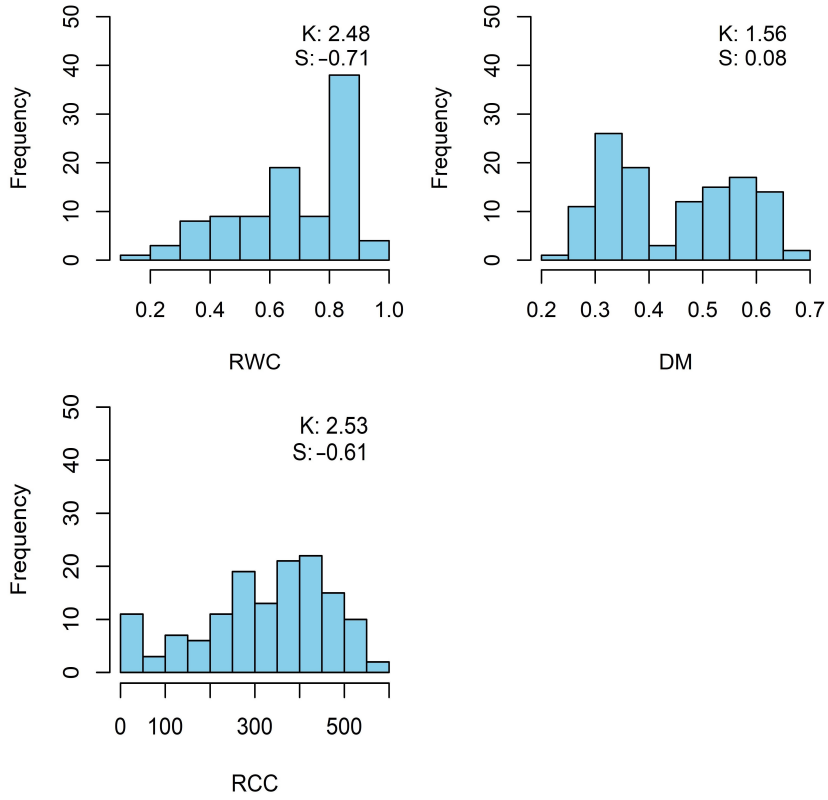


Figure 2. Frequency distributions for the ground truth data acquired for this study (relative leaf chlorophyll content, RCC; relative water content, RWC; aboveground dry matter, DM) with their respective kurtosis (K) and skewness (S) values.

2.3. Satellite Images and Processing

In this work, reflectance data were extracted from the images acquired using SuperDove satellites of the company Planet Labs PBC (hereafter Planet, San Francisco, CA, USA). The SuperDove images had a detailed spatial resolution (3 m) and 8 spectral bands (SDBs) from the visible region to the near-infrared region; Table 2 reports the spectral characteristics of the SuperDove images. The images were downloaded using the online "Planet Explorer" tool as TIFFs corrected for surface reflectance (Planet, 2024). The satellite images used to extract reflectance data were those acquired from the satellite on the same date as wheat data acquisition [45]. Reflectance data of the pixel coincident with each sampling point were extracted by using the packages "sf" and "terra" within R environment (RStudio for Windows, Version 2024.04.2 + 764, PBC, Boston, MA, USA) [46,47].

Table 2. Spectral resolution (nm) of eight bands of SuperDove images [44].

Band		Wavelength (fwhm)
B1	Coastal Blue	443 (20)
B2	Blue	490 (50)
B3	Green I	531 (36)
B4	Green	565 (36)
B5	Yellow	610 (20)
B6	Red	665 (31)
B7	Red Edge	705 (15)
B8	Nir	865 (40)

The vegetation indices calculated and used in this study are reported in Table 3. These indices were chosen for their proven sensitivity to crop parameters (e.g., chlorophyll content) using specific wavelengths in the visible and near-infrared regions. For instance, indices such as TCARI, MCARI, and their integrated forms (TCARI/OSAVI and MCARI/OSAVI) were designed to minimize the effects of disturbances such as leaf area (LAI), soil reflectance, and non-photosynthetic materials, improving the reliability of the estimates. In addition, experimental validation demonstrated a high correlation with the chlorophyll content, confirming their effectiveness at both the leaf and canopy scales, making them ideal tools for precision agriculture applications [48]. Other indices were also selected for their ability to predict plant water content, photosynthesis efficiency, and crop productivity. As demonstrated by Zhang et al., diffuse reflectance in the VIS/NIR range (400–1100 nm) can effectively detect water content while providing information on other biochemical parameters such as chlorophyll and nitrogen. This approach, combined with robust calibration models and optimized wavelength selection, ensures accurate and reliable water content estimation, supporting the development of efficient and non-invasive monitoring tools [49]. Moreover, to provide a preliminary description of field conditions before sowing (bare soil), the Brightness Index (BI) [50] was calculated by using a satellite image of the investigated area (Planet Imagery© 2023) as follows (3):

$$BI = \sqrt{RED^2 + NIR^2} \quad (3)$$

Table 3. Vegetation indices and their formulas used in this study.

Vegetation Index	Formula	Reference
Enhanced Vegetation Index (EVI)	$2 \times (\text{Nir} - \text{Red})/((\text{Nir} + 6\text{red} - 7.5\text{blue}) + 1)$	[51]
Green Normalized Vegetation Index (GNDVI)	$\text{Nir} - \text{Green}/\text{Nir} + \text{Geen}$	[52]
Modified Soil-Adjusted Vegetation Index 1 (MSAVI1)	$(1 + \text{L}) \times ((\text{Nir} - \text{Red})/(\text{Nir} + \text{Red} + \text{L}))$	[53]
Modified Soil-Adjusted Vegetation Index 2 (MSAVI2)	$\text{Nir} + 0.5 - \sqrt{(2 \times \text{Nir} + 1)^2 - 8 \times (\text{Nir} - \text{Red})}/2$	[54]
Optimized Soil-Adjusted Vegetation Index (OSAVI)	$(1 + 0.16) \times (\text{Nir} - \text{Red})/(\text{Nir} + \text{Red} + 0.1)$	[54]
Normalized Difference Red Edge (NDRE)	$(\text{Nir} - \text{Red Edge})/(\text{Nir} + \text{Red Edge})$	[55]
Normalized Difference Vegetation Index (NDVI)	$(\text{Nir} - \text{Red})/(\text{Nir} + \text{Red})$	[56]
Soil-Adjusted Vegetation Index (SAVI)	$(\text{Nir} - \text{Red})/(\text{Nir} + \text{Red} + 0.5) \times (1 + 0.5)$	[57]
Simplified Canopy Chlorophyll Content Index (SCCCI)	NDRE/NDVI	[58]
Modified Chlorophyll Absorption Reflectance Index 1 (MCARI1)	$(\text{Red Edge} - \text{Red} - 0.2 \times (\text{Red Edge} - \text{Green})) \times (\text{Red Edge}/\text{Red})$	[59]
Modified Chlorophyll Absorption Reflectance Index 2 (MCARI2)	$1.5 \times (2.5(\text{Nir} - \text{Red Edge}) - 1.3 \times (\text{Nir} - \text{Green}))$ $(2 \times (\text{Nir} + 1)\ 2 - (6 \times \text{Nir} - 5 \times (\text{Red})2) - 0.5$	[59]
Transformed Chlorophyll Absorption Reflectance Index (TCARI)	$3 \times ((\text{Red Edge} - \text{Red}) - 0.2 \times (\text{Red Edge} - \text{Green}) \times (\text{Red Edge}/\text{Red}))$	[60]
Red Edge Chlorophyll Index 1 (Cl1)	$\text{Nir}/\text{Red Edge} - 1$	[61]
Red Edge Chlorophyll Index 2 (Cl2)	$\text{Red Edge}/\text{Green} - 1$	[62]
Structure-Insensitive Pigment Index (SIPI)	$(\text{Nir} - \text{Blue})/(\text{Nir} + \text{Red})$	[63]
TCARI/OSAVI	TCARI/OSAVI	[59]
MCARI/OSAVI	MCARI/OSAVI	[59]

2.4. Modeling

To predict each wheat parameter (target variables) as a function of SuperDove reflectance products, two independent machine learning analyses were conducted with two different sets of predictors: the reflectance data of the spectral bands and the calculated vegetation indices. Figure 3 reports the framework of this study. The modeling approach compared and used in this study involved different machine learning algorithms: random forest (RF), extreme gradient boosting (XGB), and support vector regressor (SVR). RF is a powerful nonparametric method for estimating a regression function developed by Breiman [64] that has been widely used in agriculture for predicting several crop parameters [65–67]. It involves building multiple randomized decision trees during training and outputting the average of the individual trees' predictions. This ensemble method enhances the stability and accuracy of the predictive model by reducing overfitting, typical of single decision trees. The RF algorithm resamples the training data and uses random subsets of features for each tree split, producing variability that enhances prediction robustness. Each tree in the forest provides a prediction, and the final output is the average of all individual tree predictions [68]. XGB is a machine learning algorithm that has been widely adopted in various fields, including agriculture, due to its high performance and efficiency [69,70]; it is particularly known for its ability to handle large datasets and complex patterns with high accuracy [40]. XGB works by constructing a set of decision trees sequentially, where each tree tries to correct the errors of its predecessor. One of the key points of XGB is the use of gradient boosting, which involves calculating the gradient of the loss function for the model predictions and using this gradient to update the model parameters in a way that reduces loss [40]. Support vector regression (SVR) is another machine learning technique used to predict continuous outcomes by finding a function that approximates the relationship between input variables and a target variable. It has been adopted in agriculture to solve several regression tasks (e.g., nitrogen content) [41]. The SVR algorithm aims to find a hyperplane that best fits the data while allowing for a margin of tolerance between the predicted and actual values. An important aspect of SVR is the identification of support vectors, i.e., the data points that are closest to the hyperplane and are necessary in defining the position and orientation of the hyperplane. These support vectors help SVR achieve a balance between model complexity and prediction accuracy, ensuring that only the most informative data points influence the regression function [41]. For all of the modeling approaches utilized, the dataset was randomly split into a training set (80%) to calibrate the model and a testing set (20%) to test the performance and robustness of the model [71]. Machine learning analyses were carried out in RStudio. RF was implemented using the "caret", "ranger", and "randomForest" packages [71,72]. XGB was implemented with the packages "caret" and "xgboost" [73]. SVR was implemented within "caret" through the radial basis function kernel. For all of the calibrated models, 10-fold cross-validation was applied using the trainControl() function from the "caret" package to reduce the risk of overfitting and to fine-tune the hyperparameters. To this aim, a systematic grid search approach, using the expand.grid() function, was employed to explore different hyperparameter combinations and, during the tuning process, the performance of each model was assessed using cross-validation; thus, the best set of hyperparameters was selected based on these results [74]. Table 4 reports the machine learning algorithm and R packages used and the hyperparameters that were fine-tuned. This method involves dividing the entire dataset into ten subsets of equal size or folds. In each iteration of the cross-validation process, one-fold is held as the validation set, while the other nine folds are used to train the model. The process was repeated over multiple random seeds to ensure the robustness of model performance. The coefficient of determination (R^2, Equation (4)) and the root

mean square error (RMSE, Equation (5)) were calculated as follows for each iteration and averaged across the iterations to obtain the final evaluation metrics:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2} \quad (5)$$

where SSres is the sum of the squares of the residuals, SStot is the total sum of the squares, P is the predicted value, O is the observed value, and n is the number of observations. Graphs, images, and maps were made using RStudio, Sigmaplot (SigmaPlot, Systat Software Inc, Version 14 for Windows, San Jose, CA, USA), and QGis (QGis, version 3.34.9 for Windows).

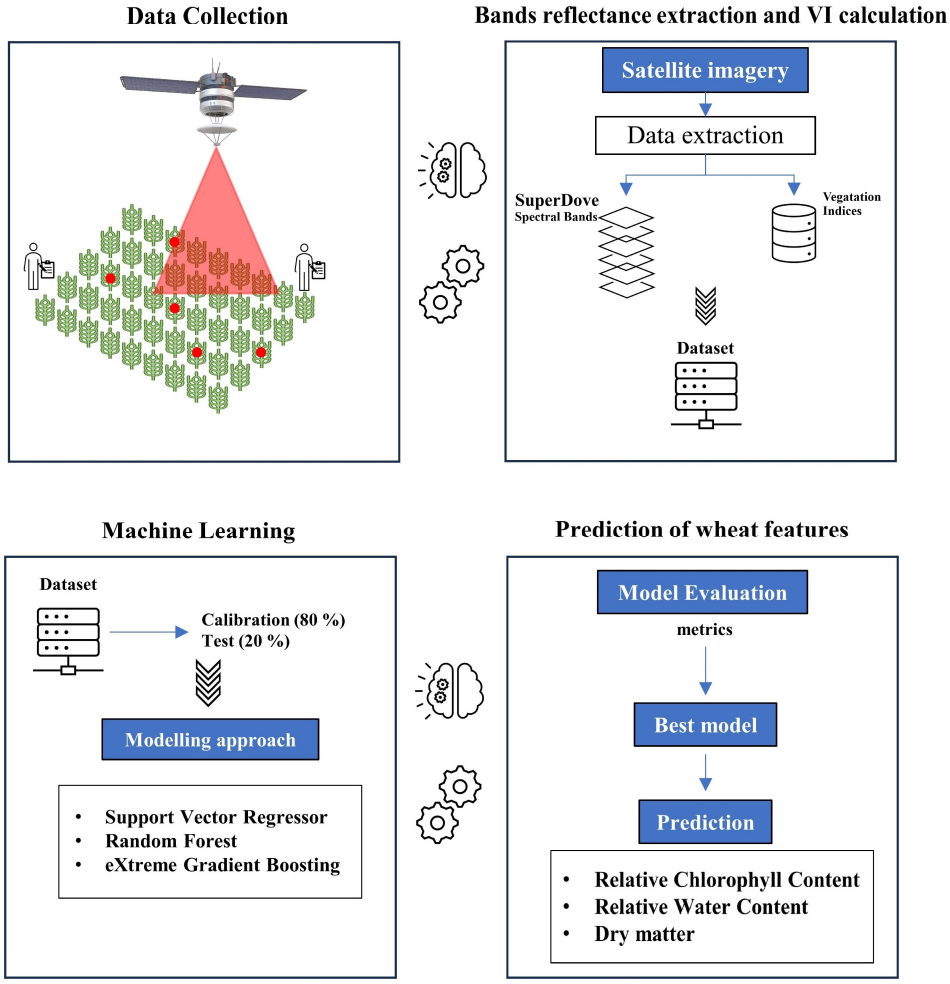


Figure 3. The framework of the study for the prediction of the relative leaf chlorophyll content, relative water content, and aboveground dry matter.

2.5. Machine Learning Inference and Interpretation

Once the best model for each target variable was individuated among the trained ones and showed good results in testing, it was applied to produce maps and graphs showing the variability of the considered parameter. Maps were produced for the date of field measurement of each target variable. To model the temporal variability of each parameter, the best model found was applied to further spectral images (according to the usefulness of the images in relation to the weather conditions, there were 3 in the case of RCC and 1 in the case of the RWC and DM) from the date of the first measurement to pre-harvest, and the average value was estimated.

Table 4. Models, RStudio packages, and hyperparameters that were fine-tuned.

Model	RStudio Package	Hyperparameters
Random forest	caret; randomForest; ranger	mtry (the number of variables to potentially split each node); splitting rule; minimum node size
Extreme gradient boosting	caret; xgboost	the maximum depth of each decision tree used in the boosting process; eta (the learning rate of the model)
Support vector regressor	caret	c (the trade-off between maximizing the margin of the decision boundary and minimizing the regression errors); sigma (in the radial basis function kernel, it determines the reach of a single training instance)

For each best model built, a machine learning interpretation technique was used to detect the driving factors of the variability. Particularly, the SHapley Additive exPlanations (SHAP) method was used to interpret the output of the model for each target. SHAP employs the "Shapley value" from cooperative game theory, which measures the average marginal contribution of a player in all potential coalitions; in the context of machine learning, it is used to explain the feature's importance to a predicted variable and then how it increased or decreased the prediction [75].

3. Results

3.1. Agro-Meteorological Conditions

During the growing season, the lowest average temperature recorded was 4.25 °C on the 30th day of the year (DOY), while the highest one was recorded on DOY 161 (28.59 °C). Figure 4A reports the daily variation in the climatic data during the wheat growing season. The peaks of ETc were recorded on DOYs 100 (9.08 mm) and 106 (8.04 mm); the total seasonal amount of ETc was 495.50 mm. The total rainfall of the season was 224.20 mm with peaks on DOYs 20 (18.40 mm), 56 (51.60 mm), and 64 (20.40 mm); after that, the rainfall was relatively reduced during the growing season (Figure 4B).

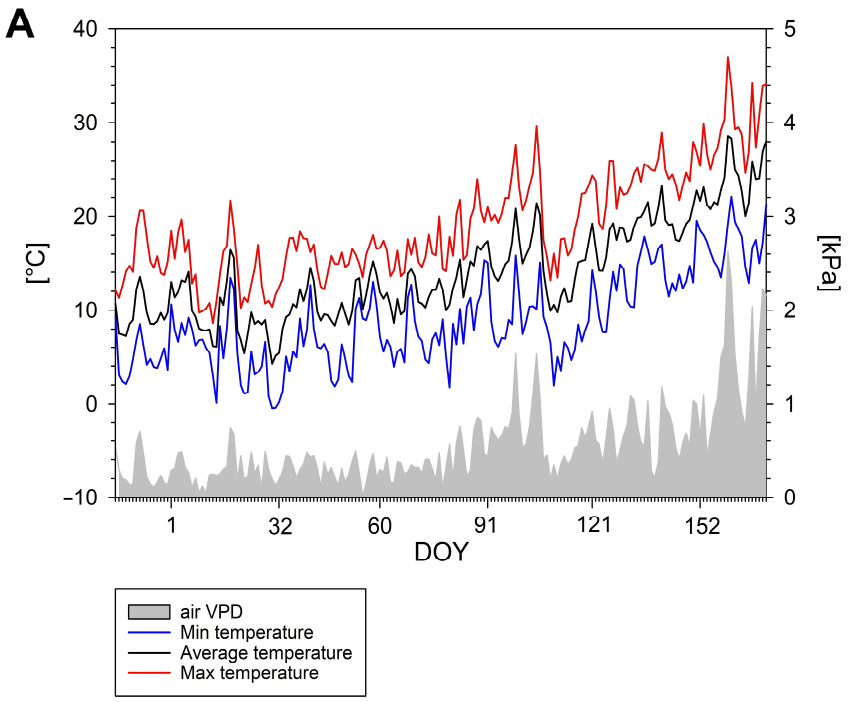


Figure 4. *Cont.*

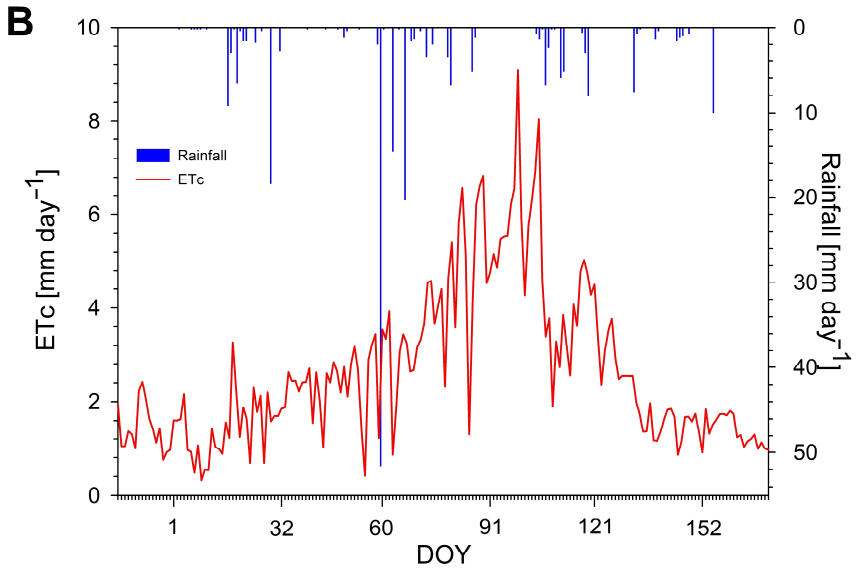


Figure 4. Variation in air vapor pressure deficit (air VPD), minimum, mean, and maximum temperatures during wheat growing season (**A**); variation in crop evapotranspiration (ETc) and rainfall during wheat growing season (**B**).

3.2. Soil Brightness Index

The analysis of the BI in the study area has shown that the pixel values ranged from a minimum of 0.137 to a maximum of 0.282, indicating a significant variation in soil brightness across the study area. The spatial distribution of BI values across the 1.46 hectares of the field was highly heterogeneous, with the southeastern part and the left border exhibiting higher values compared to the rest of the area, which was dominated by lower BI values. To further investigate the BI characteristics, the frequency distribution of pixel values was analyzed, as shown in Figure 5. The resulting distribution was right-skewed, with skewness and kurtosis values of 0.842 and 2.431, respectively, indicating a non-symmetrical and slightly flatter distribution with a longer tail towards higher values of soil brightness. Most BI values were found to cluster around the range of 0.158 to 0.165, with a mean of 0.192 and a median of 0.179, suggesting that these values were the most prevalent throughout the study area. However, it is important to note that while higher BI values were present, they were less frequent compared to the lower and mid-range values. This irregular distribution highlights the heterogeneity of soil properties within the field.

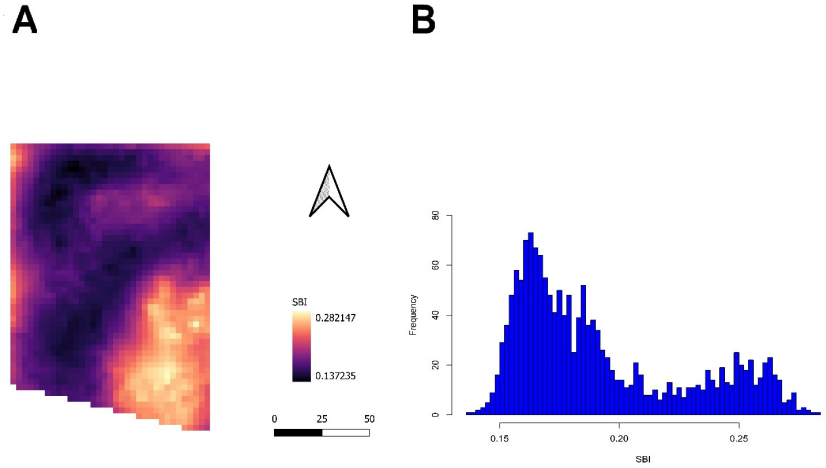


Figure 5. A map (**A**) and frequency plot of the pixel value (**B**) of the Brightness Index of the field before sowing (19 November 2023).

3.3. Models' Performance

Generally, the model that outperformed the others in the prediction of the RCC, RWC, and DM was RF. Particularly, RF achieved the best performance when trained with SDBs for predicting RWC (R^2 = 0.71) and DM (R^2 = 0.87), and with VIs for predicting RCC (R^2 = 0.75). SVR and XGB had lower performances, especially in the case of RWC prediction when using SDBs as predictors (R^2 = 0.16 and R^2 = 0.07, respectively). Table 5 reports the performance parameters of the algorithms tested for this study. A summary of the statistics of the residuals for each model, when trained with SDBs and VIs, is reported in the Supplementary Material (Table S1). The RF model consistently showed smaller residual ranges and more robust predictions for both sets of predictors. For instance, in the prediction of the RWC, the RF residuals ranged from −0.188 to 0.153 with SDBs and from −0.264 to 0.215 with VIs. On the other hand, SVR and XGB showed wider ranges (SVR: −0.551 to 0.293 with SDBs, −0.408 to 0.330 with VIs; XGB: −0.449 to 0.303 with SDBs, −0.390 to 0.253 with VIs), indicating greater sensitivity to extreme values and variability. These differences suggest that RF could handle noise and variability in the data, while SVR and XGB exhibited reduced stability under these conditions. Furthermore, the median residuals of SVR and XGB deviated more from zero than RF, particularly when trained on VIs, reflecting possible slight bias in their predictions.

Table 5. The performance parameters of the models used in this study (random forest, RF; support vector regressor, SVR; extreme gradient boosting, XGB) for the prediction of the relative leaf chlorophyll content (RCC), relative water content (RWC), and aboveground dry matter (DM) using the SuperDove spectral bands (SDBs) or the vegetation indices (VIs) as predictors; the means and standard deviations of 10 replicates are shown.

			R^2	RMSE
SDBs	RCC	RF	0.75 (±0.01)	75.92 (±2.57)
		SVR	0.72 (±0.03)	80.27 (±6.17)
		XGB	0.64 (±0.03)	86.94 (±2.64)
	DM	RF	0.87 (±0.005)	0.03 (±0.004)
		SVR	0.74 (±0.005)	0.05 (±0.005)
		XGB	0.76 (±0.001)	0.05 (±0.001)
	RWC	RF	0.71 (±0.01)	0.07 (±0.03)
		SVR	0.16 (±0.02)	0.18 (±0.004)
		XGB	0.07 (±0.04)	0.17 (±0.008)
VIs	RCC	RF	0.81 (±0.01)	39.84 (±4.65)
		SVR	0.56 (±0.01)	99.44 (±1.87)
		XGB	0.61 (±0.02)	87.64 (±2.53)
	DM	RF	0.82 (±0.01)	0.02 (±0.003)
		SVR	0.75 (±0.01)	0.06 (±0.002)
		XGB	0.77 (±0.02)	0.05 (±0.003)
	RWC	RF	0.61 (±0.03)	0.11 (±0.005)
		SVR	0.22 (±0.05)	0.18 (±0.006)
		XGB	0.36 (±0.004)	0.14 (±0.001)

Figure 6 shows SHAP value plots for the best model found for each target considered in this work. For the RF-based model, with VIs as predictors, GNDVI, Cl1, and NDRE were the most influential in the prediction of the RCC, while VIs like MCARI, MCARI/OSAVI, and TCARI had a weak influence on RCC prediction. For the prediction of the DM, the

RF-based model with SDBs as predictors, namely SDBs 5, 6, and 8 (yellow, red, and nir, respectively), had the highest influence in the prediction; in particular, SDB 5 showed an important impact, with higher values positively contributing to the predictions. For the RF-based model for the prediction of the RWC with SDBs, the features with the highest impact in predicting the RWC were SDBs 5, 8, and 6 (yellow, nir, and red, respectively).

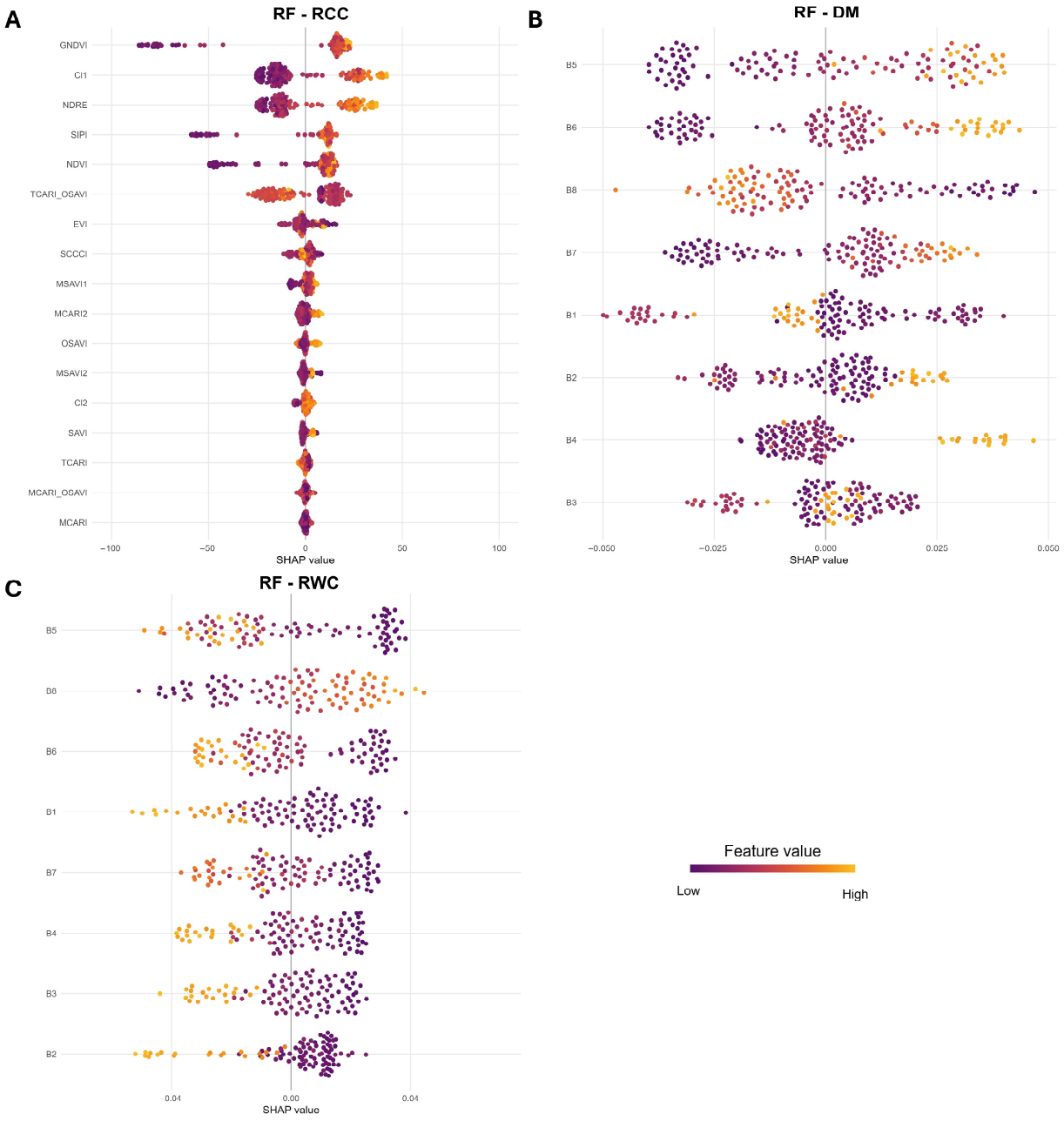


Figure 6. SHAP summary plots for the random forest model for the prediction of the relative leaf chlorophyll content (RCC, **A**); aboveground dry matter (DM, **B**); and relative water content (RWC, **C**). The graphs report the distribution of SHAP values for each feature, with the dots' colors changing depending on the features' values. A positive SHAP value indicates a positive impact on the model, and a negative SHAP value indicates a negative impact on the model.

3.4. Predicted Parameters

Figure 7 reports the trends in the RCC, DM, and RWC during the investigated season. Figure 8 shows the RCC prediction maps obtained by applying the RF-VI-based model; in the first predictive map of the variable, it is already possible to observe a reduced

concentration of leaf chlorophyll in the lowest right-hand part of the field. This condition also occurred over the subsequent three DOYs (66, 82, and 96). The variability within the field was no longer observable over the last three DOYs (117, 124, and 130). Nevertheless, the map of the average RCC over the examined date of measurements shows that the predicted RCC in the southeastern part of the plot was lower. Figure 9 illustrates the spatial distribution of DM for all six DOYs where field measurements were taken, showing little spatial variability (Figure 9). The map of the average DM on the measurement date examined shows high spatial uniformity. Figure 10 shows the spatial representation of the predicted RWC; in all of the maps in the figure, high uniformity is observed except for the map at DOY 82, which seems to present, as for the RCC variable, an area to the southeast of the field with slightly lower values. Nevertheless, the average of the maps shows high spatial uniformity.

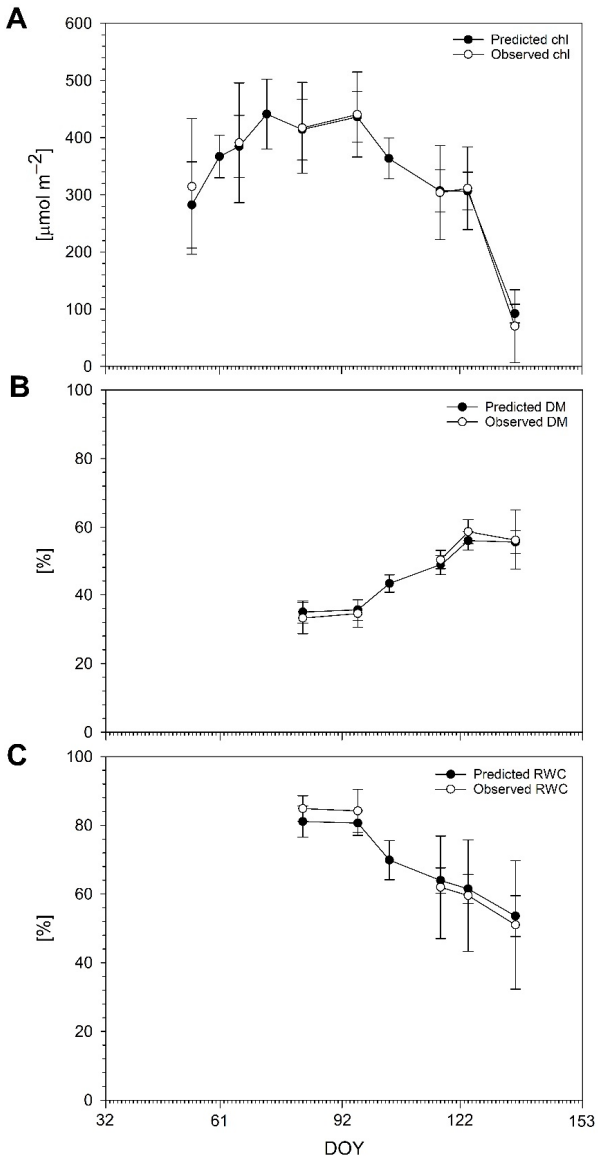


Figure 7. The trends in the mean and standard deviation of the observed and predicted variables: (**A**) the relative chlorophyll content, (**B**) aboveground dry matter, and (**C**) relative water content. The observed and predicted values are compared on the same DOYs for each parameter. Moreover, predictions were extended to additional DOYs, namely 61, 73, and 104 in the case of RCC and 104 in the case of DM and RWC.

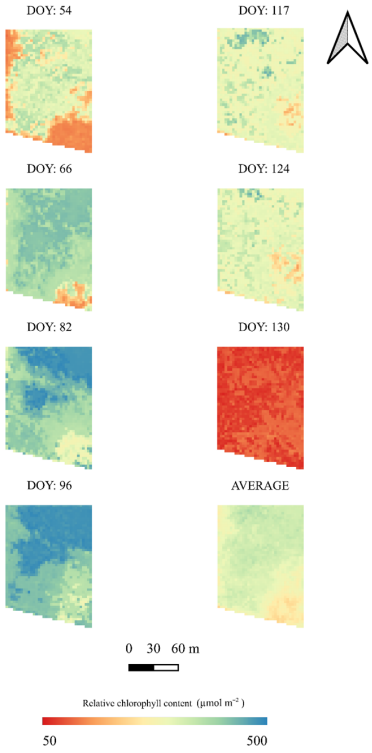


Figure 8. Predictive maps of the relative leaf chlorophyll content for the DOY (day of the year) of ground measurements, obtained by applying the random forest model trained with vegetation indices.

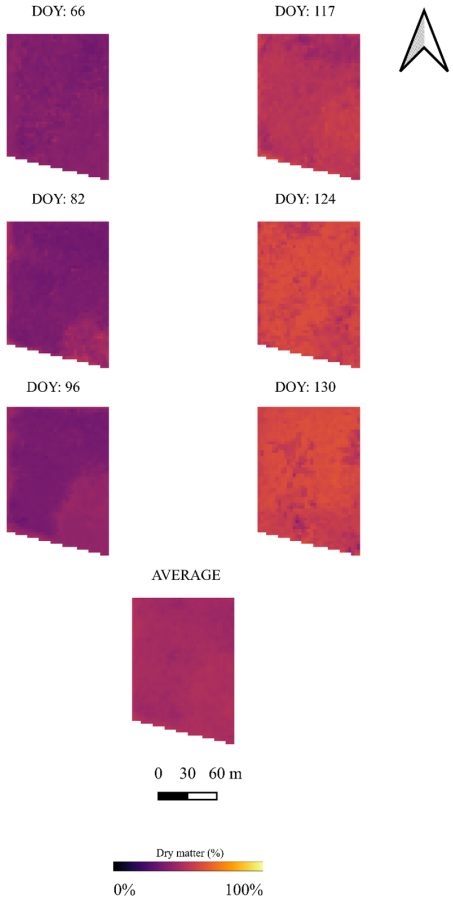


Figure 9. Predictive maps of the aboveground dry matter for the DOY (day of the year) of ground measurements, obtained by applying the random forest model trained with SuperDove spectral bands.

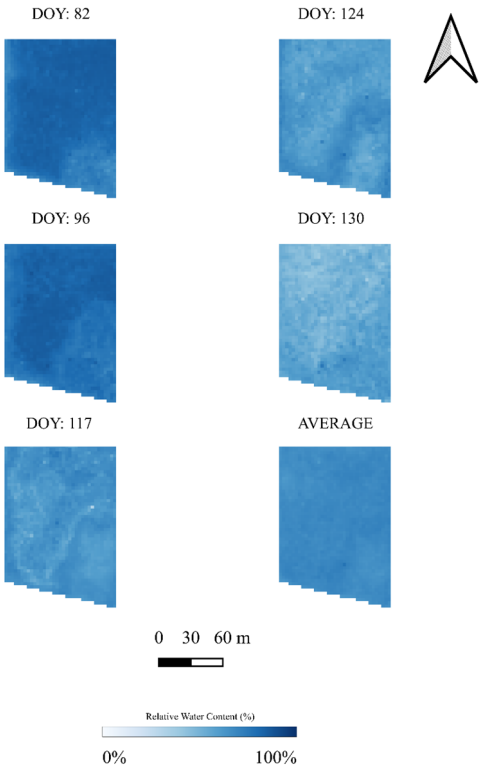


Figure 10. Predictive maps of the relative water content for the DOY (day of the year) of ground measurements, obtained by applying the random forest model trained with SuperDove spectral bands.

4. Discussion

Agriculture is increasingly stressed by climate change, which impacts crop yields and quality through altered temperatures, precipitation patterns, and more frequent extreme weather events [76]. Improved crop monitoring technologies are essential to address these challenges. Remote sensing technologies can improve the accuracy of monitoring crop status and development [77]. The predictive maps of the RWC and RCC can be used to optimize irrigation and nitrogen fertilization planning, thus providing a better and more efficient monitoring of crop performance, with potential positive impacts on the environment [78]. In addition, machine learning models integrating VIs, soil properties, and soil electrical conductivity have proved to be useful in defining irrigation management zones [78]. Monitoring maps can provide a practical tool for the early detection of different stresses; they can be useful for efficient irrigation management, especially at a time when drought events are becoming more extreme and frequent [79].

4.1. Platforms Comparison

This study investigated the use of SuperDove satellite imagery and various models to predict some parameters of winter wheat. To the best of the authors' knowledge, this is the first study investigating the use of SuperDove fine-scale imagery for predicting the RCC, RWC, and DM of winter wheat in Mediterranean environments. SuperDove images have a spatial resolution of 3 m and a daily revisit frequency, potentially allowing for detailed and timely monitoring of crop conditions [45]. In contrast, other satellites, such as Sentinel-2 and Landsat 8 and 9, provide multispectral images with lower spatial details and revisit times. Thus, the use of SuperDove images could allow for more detailed monitoring of crop variability at the field level, which is a key aspect in precision agriculture [80]. In addition, the high temporal frequency of SuperDove satellites compared to other platforms could reduce the impact of cloud cover, ensuring more consistent data acquisition.

4.2. Models Performance

In model comparison, RF outperformed the other models. RF probably performed better because of its ensemble learning nature in which the aggregation of multiple decision trees reduces variance and mitigates overfitting, resulting in more stable and robust predictions; in addition, its inherent structure allows for the identification of complex, nonlinear relationships between predictors and the target variable [81]. The lower performance of SVR and XGB, compared to RF, can be attributed to factors related to the characteristics of the algorithm and/or the nature of the datasets. SVR focuses on a limited subset of data points (support vectors) to define the regression model, making it more sensitive to noise and less able to exploit all of the information in the dataset. This feature could be particularly limiting in scenarios where the data distribution is not uniform or where predictors contain scattered patterns that are better captured by ensemble-based approaches such as RF [82,83]. In the case of XGB, its reliance on sequentially constructed decision trees may have been limited by the relatively small size and variability of the dataset. Gradient boosting algorithms typically work best with large datasets, where models can be iteratively refined over multiple boosting rounds. The limited datasets in this study may have contained noise that affected the boosting process, causing overfitting or unstable predictions [84,85].

4.2.1. Relative Chlorophyll Content Prediction

The RF model was found to be the best-performing model in the prediction of the RCC, demonstrating the model's effectiveness in capturing wheat chlorophyll variability. In this case, RF had better results when using the VIs as predictors. In a study conducted by Shah et al. [67], RF had a significantly higher accuracy in chlorophyll content prediction when trained with VIs rather than when using the entire range of spectral bands. Wang et al. [86] predicted the wheat chlorophyll content using drone-derived VIs through different learning-based models, and the authors found that the Ridge and SVM models outperformed the other algorithms, including RF. In this work, the higher performance of RF trained with VIs rather than SDBs could be explained by the fact that the VIs—especially those designed for chlorophyll monitoring—provide the model with more informative inputs than single spectral bands for RCC prediction. Monitoring the chlorophyll content could help in fertilization management since it can predict the nitrogen needs of wheat and serve as a vital indicator of the photosynthetic apparatus status [87]. The temporal trend in the RCC predicted by the RF model showed a gradual decline as the crop approached senescence; this decline is consistent with the natural reduction in chlorophyll as the crop approaches maturity.

4.2.2. Dry Matter Prediction

The RF model trained with SDBs achieved an R^2 of 0.87 for the prediction of DM, making it the best result of this study. This high level of accuracy highlights the usefulness of the machine learning method combined with satellite data for estimating biomass-related variables such as DM, which is a key parameter for crop growth and yield potential assessment. The results of the RF-based model are in line with the literature; for instance, Jiang et al. [65] obtained interesting results in the prediction of wheat DM using satellite imagery (Sentinel-2) and RF, which outclassed the other tested algorithms. The predicted DM values showed a steady increase during the growing season, reflecting the expected pattern of biomass accumulation in wheat. Crop productivity can be evaluated in terms of DM accumulation, and specifically, grain production needs higher DM production during the vegetative growing phase [88]. Similarly, enhanced post-anthesis dry matter accumulation resulted in a higher thousand kernels weight in winter wheat [89]. Knowing

this trend is relevant to inform management practices such as fertilization, which assumes a central role in achieving a successful yield, especially in winter wheat [90,91]. More specifically, an adequate reduction in nitrogen supply alleviates the combined effect of drought and heat stresses on the aboveground dry mass and green leaf area.

4.2.3. Relative Water Content Prediction

Also, in the case of the prediction of the RWC, the RF-based model outperformed the other ones when using SDBs as predictors. Several studies have investigated the use of satellite remote sensing in the monitoring of the canopy water content [92,93], but this research is the first to integrate machine learning and detailed spatial imagery to predict the RWC specifically. The RWC is a reliable indicator of wheat water status and is positively correlated with grain yield, especially when plants have a higher RWC during the anthesis and mid-grain filling phases [44]. Mokhtari et al. [94] reported that the RWC exhibited a significant positive relationship with the yield stability index, which supports the selection of yield-stable genotypes based on a higher RWC under drought conditions. The predicted RWC showed a decreasing trend during the growing season, reflecting the natural progression of water content reduction as the crop ripens [95]. This trend is particularly important for irrigation management as it allows farmers to adjust water applications according to wheat needs, thereby improving water use efficiency and reducing the risk of over- or under-irrigation in specific areas where wheat is grown under irrigation conditions [96]. Using the predicted RWC together with the appropriate water-saving genotype recommended for every agroecological zone will be of great interest in future irrigation scheduling. The model evaluation results suggest that the choice of predictors in crop remote sensing applications can significantly influence prediction accuracy. While some studies have reported that the use of spectral bands rather than vegetation indices improves model performance [97,98], this work suggests that it cannot be generalized that machine learning models consistently perform better in agricultural remote sensing applications when trained with spectral band reflectance, and their performance may depend on several factors, such as the crop, target parameter, agro-climatic conditions, site-specific traits, or others; further studies are needed to clarify this aspect.

4.2.4. SHAP Analysis

A SHAP analysis provided insights into model interpretation and explainability. The SHAP analysis showed that the SDB 5 band was the most influential in predicting DM. The contribution of SDBs 6 and 8 could have further improved the sensitivity of the model, especially when used for data from the period during the transition from vegetative growth to maturity, where biomass changes are more marked; visible and nir reflectance regions are closely related to vegetation density and biomass [99,100]. The SHAP values identified GNDVI, Cl1, and NDRE as the most influential indices for predicting the RCC, in line with their known role in monitoring chlorophyll levels [58,101]. High values for these features tend to be associated with positive SHAP values, suggesting that high values of these variables contribute positively to RCC prediction. These indices are particularly effective in detecting slight changes in the chlorophyll content, which are useful for early crop management intervention [102]. SDBs 5, 6, and 8 were the most influential in the SHAP analysis for RWC prediction. In particular, SDB 5 shows that high values are associated with negative SHAP values, while for SDB 8, high values are associated with positive SHAP values; in the case of SDB 6, the impact is more distributed between positive and negative values. Nir and red spectral regions are well known to be highly related to plant water status [103,104].

5. Conclusions

This is the first study on the integration of predictive algorithms and SuperDove high-resolution imagery for the prediction of the relative leaf chlorophyll content, relative water content, and aboveground dry matter of winter wheat. This represents a step forward in supporting sustainable agricultural practices by offering precision in crop monitoring and management. Combining high-resolution satellite imagery with machine learning techniques can provide pragmatic solutions to increase resource efficiency in agriculture by optimizing crop productivity and minimizing inputs such as water and fertilizers. In terms of machine learning performance, this study highlights the differences between vegetation indices and spectral bands as predictors. Random forest proved to be the best-performing algorithm depending on the predictors used (vegetation indices and spectral bands). However, some limitations should be addressed in future research. While SuperDove's spatial resolution can be advantageous for capturing fine-scale heterogeneity within the wheat field, its spectral resolution is relatively limited compared to platforms like Sentinel-2, which includes specific red-edge bands designed for vegetation monitoring and short-wave infrared bands. While the Brightness Index provided a preliminary description of field conditions prior to sowing, future studies should incorporate detailed soil analyses to better explain the relationship between soil properties and crop variability. The use of specific soil information would provide a more complete overview to link soil heterogeneity to the predicted crop parameters. This integration could significantly improve the accuracy and applicability of remote sensing and machine learning approaches in precision agriculture. Future work could benefit from the integration of SuperDove data with that of other satellite platforms such as Sentinel-2; this integration could improve the accuracy of monitoring wheat parameters by combining the frequent acquisition and fine spatial detail of SuperDove with the spectral diversity of Sentinel-2, including red edge bands and shortwave infrared bands, which are useful for assessing crop health and water stress. In addition, it is recommended to extend the analyses to a larger regional scale to assess the applicability and effectiveness of the developed models under different agroecological conditions and farming practices. This approach would validate the robustness of the proposed methods and facilitate their adaptation to heterogeneous agricultural contexts. Aligned with the principles of sustainable agriculture, this study provides a pathway to reduce environmental impact and optimize agricultural productivity. Future advances in satellite technology and machine learning will further enhance the capabilities of this approach, paving the way for precision farming solutions that are not only accurate and efficient, but also accessible to a wider range of stakeholders worldwide.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/agronomy15010241/s1, Table S1. A summary of the statistics of the residuals for the random forest (RF), support vector regression (SVR), and extreme gradient boosting (XGB) models trained with spectral bands (SDBs) and vegetation indices (VIs) for predicting the relative chlorophyll content (RCC), dry matter (DM), and relative water content (RWC). The table reports the minimum (Min), median, mean, third quartile (3rd Qu.), and maximum (Max) residual values.

Author Contributions: Conceptualization, G.D.C., V.G., M.M. and S.P.G.; Data Curation, G.D.C., L.C., F.A., S.R. and S.P.G.; Formal Analysis, G.D.C., S.P.G., A.M.S. and A.T.; Investigation, G.D.C., L.C., F.A., N.S. and S.P.G.; Methodology, G.D.C., V.G., L.C., F.A., S.T., M.M. and S.P.G.; Resources, N.S.; Software, L.C., F.A. and S.P.G.; Supervision, S.T., M.M. and S.P.G.; Visualization, L.C. and S.P.G.; Writing—Original Draft, G.D.C., L.C., F.A. and S.P.G.; Writing—Review and Editing, V.G., A.M.S., A.T., S.R., S.T., M.M., N.S. and S.P.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors thank Planet Labs PBC for providing images as part of Planet's Education and Research Program. The authors thank Beniamino Leoni for his kind technical support.

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

References

1. Hunter, M.C.; Smith, R.G.; Schipanski, M.E.; Atwood, L.W.; Mortensen, D.A. Mortensen, Agriculture in 2050: Recalibrating Targets for Sustainable Intensification. *BioScience* **2017**, *67*, 386–391. [CrossRef]
2. Neupane, J.; Wenxuan, G. Agronomic Basis and Strategies for Precision Water Management: A Review. *Agronomy* **2019**, *9*, 87. [CrossRef]
3. Barca, E.; De Benedetto, D.; Stellacci, A.M. Optimization of sampling design for soil total organic carbon assessment in the precision agriculture framework: Impact of different variogram models and potentiality of ground penetrating radar (GPR) covariate information. *Comput. Electron. Agric.* **2024**, *226*, 109470. [CrossRef]
4. Singh, R.; Singh, G.S. Traditional agriculture: A climate-smart approach for sustainable food production. *Energy Ecol. Environ.* **2017**, *2*, 296–316. [CrossRef]
5. Wezel, A.; Soboksa, G.; McClelland, S.; Delespesse, F.; Boissau, A. The blurred boundaries of ecological, sustainable, and agroecological intensification: A review. *Agron. Sustain. Dev.* **2015**, *35*, 1283–1295. [CrossRef]
6. Tilman, D.; Balzer, C.; Hill, J.; Befort, B.L. Befort. Global Food Demand and the Sustainable Intensification of Agriculture. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20260–20264. [CrossRef]
7. Lago-Olveira, S.; Rebolledo-Leiva, R.; Garofalo, P.; Moreira, M.T.; González-García, S. Environmental and Economic Benefits of Wheat and Chickpea Crop Rotation in the Mediterranean Region of Apulia (Italy). *Sci. Total Environ.* **2023**, *896*, 165124. [CrossRef]
8. Erenstein, O.; Jaleta, M.; Mottaleb, K.A.; Sonder, K.; Donovan, J.; Braun, H.-J. Global trends in wheat production, consumption and trade. In *Wheat Improvement*; Springer International Publishing: Cham, Switzerland, **2022**; pp. 47–66. [CrossRef]
9. Poole, N.; Donovan, J.; Erenstein, O. Viewpoint: Agri-nutrition research: Revisiting the contribution of maize and wheat to human nutrition and health. *Food Policy* **2021**, *100*, 101976. [CrossRef] [PubMed]
10. Luthria, D.L.; Lu, Y.; John, K.M.M. Bioactive phytochemicals in wheat: Extraction, analysis, processing, and functional properties. *J. Funct. Foods* **2015**, *18*, 910–925. [CrossRef]
11. Zingale, S.; Spina, A.; Ingrao, C.; Fallico, B.; Timpanaro, G.; Anastasi, U.; Guarnaccia, P. Factors affecting the nutritional, health, and technological quality of durum wheat for pasta-making: A systematic literature review. *Plants* **2023**, *12*, 530. [CrossRef] [PubMed]
12. EUROSTAT. *Wheat and Spelt by Area, Production and Humidity*; EUROSTAT: Luxembourg, **2023**.
13. Le Gouis, J.; Oury, F.-X.; Charmet, G. How changes in climate and agricultural practices influenced wheat production in Western Europe. *J. Cereal Sci.* **2020**, *93*, 102960. [CrossRef]
14. Ru, C.; Hu, X.; Chen, D.; Wang, W.; Zhen, J.; Song, T. Individual and combined effects of heat and drought and subsequent recovery on winter wheat (*Triticum aestivum* L.) photosynthesis, nitrogen metabolism, cell osmoregulation, and yield formation. *Plant Physiol. Biochem.* **2023**, *196*, 222–235. [CrossRef]
15. Riedesel, L.; Möller, M.; Piepho, H.-P.; Rentel, D.; Lichthardt, C.; Golla, B.; Kautz, T.; Feike, T. Site conditions determine heat and drought induced yield losses in wheat and rye in Germany. *Environ. Res. Lett.* **2024**, *19*, 034024. [CrossRef]
16. Muhammed, S.E.; Coleman, K.; Wu, L.; Bell, V.A.; Davies, J.A.C.; Quinton, J.N.; Carnell, E.J.; Tomlinson, S.J.; Dore, A.J.; Dragosits, U.; et al. Impact of two centuries of intensive agriculture on soil carbon, nitrogen and phosphorus cycling in the UK. *Sci. Total Environ.* **2018**, *634*, 1486–1504. [CrossRef] [PubMed]
17. Rose, M.; Pahlmann, I.; Kage, H. Modified crop rotations for a sustainable intensification? A case study in a high-yielding environment with recurrent nitrogen surplus. *Eur. J. Agron.* **2023**, *142*, 126644. [CrossRef]
18. Colovic, M.; Stellacci, A.M.; Mzid, N.; Di Venosa, M.; Todorovic, M.; Cantore, V.; Albrizio, R. Comparative Performance of Aerial RGB vs. Ground Hyperspectral Indices for Evaluating Water and Nitrogen Status in Sweet Maize. *Agronomy* **2024**, *14*, 562. [CrossRef]
19. Fox, R.H.; Piekielek, W.P.; Macneal, K.M. Using a chlorophyll meter to predict nitrogen fertilizer needs of winter wheat. *Commun. Soil. Sci. Plant Anal.* **1994**, *25*, 171–181. [CrossRef]
20. Liu, T.; Zhang, H.; Wang, Z.Y.; He, C.; Zhang, Q.G.; Jiao, Y.M. Estimation of the leaf area index and chlorophyll content of wheat using UAV multi-spectrum images. *J. Trans. Chin. Soc. Agric. Eng.* **2021**, *37*, 65–72. [CrossRef]

21. Qiang, W.; Zhang, Y.; Zhao, Z.; Xie, M.; Hou, D. Estimation of Relative Chlorophyll Content in Spring Wheat Based on Multi-Temporal UAV Remote Sensing. *Agronomy* **2023**, *13*, 211. [CrossRef]
22. Yang, B.; Wen, H.; Wang, S.; Zhang, J.; Wang, Y.; Zhang, T.; Yuan, K.; Lu, L.; Liu, Y.; Xue, Q.; et al. Enhancing Drought Resistance and Yield of Wheat through Inoculation with Streptomyces pactum Act12 in Drought Field Environments. *Agronomy* **2024**, *14*, 692. [CrossRef]
23. Memon, H.M.U.; Bux, H.; Sial, M.A.; Hussain, N.; Soomro, S.A.; Mujtaba, S.; Bughio, F.A. Genetic screening of wheat using multivariate analysis under moisture stress. *Pak. J. Bot.* **2024**, *56*, 1309–1320. [CrossRef]
24. Drees, L.; Demie, D.T.; Paul, M.R.; Leonhardt, J.; Seidel, S.J.; Döring, T.F.; Roscher, R. Data-driven crop growth simulation on time-varying generated images using multi-conditional generative adversarial networks. *Plant Methods* **2024**, *20*, 93. [CrossRef] [PubMed]
25. Garofalo, S.P.; Giannico, V.; Lorente, B.; Vivaldi, G.A.; Jose, A. Predicting carob tree physiological parameters under different irrigation systems using Random Forest and Planet satellite images. *Front. Plant Sci.* **2024**, *15*, 1302435. [CrossRef] [PubMed]
26. Stamford, J.D.; Vialet-Chabrand, S.; Cameron, I.; Lawson, T. Development of an accurate low cost NDVI imaging system for assessing plant health. *Plant Methods* **2023**, *19*, 9. [CrossRef] [PubMed]
27. Giannico, V.; Garofalo, S.P.; Brillante, L.; Sciusco, P.; Elia, M.; Lopriore, G.; Camposeo, S.; Lafortezza, R.; Sanesi, G.; Vivaldi, G.A. Temporal Vine Water Status Modeling Through Machine Learning Ensemble Technique and Sentinel-2 Multispectral Images Under Semi-Arid Conditions. *Remote Sens.* **2024**, *16*, 4784. [CrossRef]
28. Pan, H.; Chen, Z.; Ren, J.; Li, H.; Wu, S. Modeling Winter Wheat Leaf Area Index and Canopy Water Content With Three Different Approaches Using Sentinel-2 Multispectral Instrument Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 482–492. [CrossRef]
29. Ogungbuyi, M.G.; Guerschman, J.; Fischer, A.M.; Crabbe, R.A.; Ara, I.; Mohammed, C.; Scarth, P.; Tickle, P.; Whitehead, J.; Harrison, M.T. Improvement of pasture biomass modelling using high-resolution satellite imagery and machine learning. *J. Environ. Manag.* **2024**, *356*, 120564. [CrossRef]
30. Ali, A.M.; Darvishzadeh, R.; Skidmore, A.K. Retrieval of specific leaf area from landsat-8 surface reflectance data using statistical and physical models. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 3529–3536. [CrossRef]
31. Gawdiya, S.; Kumar, D.; Sharma, R.K.; Das, P.; Choudhary, M.; Matter, M.A. Field scale wheat yield prediction using ensemble machine learning techniques. *Smart Agric. Technol.* **2024**, *9*, 100543. [CrossRef]
32. Verrelst, J.; Munoz, J.; Alonso, L.; Delegido, J.; Rivera, J.P.; Camps-Valls, G.; Moreno, J. Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and-3. *Remote Sens. Environ.* **2012**, *118*, 127–139. [CrossRef]
33. Wang, J.; Xiao, X.; Bajgain, R.; Starks, P.; Steiner, J.; Doughty, R.B.; Chang, Q. Estimating leaf area index and aboveground biomass of grazing pastures using Sentinel-1, Sentinel-2 and Landsat images. *ISPRS J. Photogramm. Remote Sens.* **2019**, *154*, 189–201. [CrossRef]
34. Fei, S.; Li, L.; Han, Z.; Chen, Z.; Xiao, Y. Combining novel feature selection strategy and hyperspectral vegetation indices to predict crop yield. *Plant Methods* **2022**, *18*, 119. [CrossRef] [PubMed]
35. Van Klompenburg, T.; Kassahun, A.; Catal, C. Crop yield prediction using machine learning: A systematic literature review. *Comput. Electron. Agric.* **2020**, *177*, 105709. [CrossRef]
36. Vapnik, V.N. *The Nature of Statistical Learning Theory*; Springer: New York, NY, USA, **1995**; p. 188. [CrossRef]
37. Yadav, V.P.; Prasad, R.; Bala, R.; Srivastava, P.K. Synergy of vegetation and soil microwave scattering model for leaf area index retrieval using C-band sentinel-1Asatellite data. *IEEE Geosci. Remote Sens.* **2020**, *19*, 4004605. [CrossRef]
38. Yadav, V.P.; Prasad, R.; Bala, R.; Vishwakarma, A.K. An improved inversion algorithm for spatio-temporal retrieval of soil moisture through modified water cloud model using C-band sentinel-1A SAR data. *Comput. Electron. Agric.* **2020**, *173*, 105447. [CrossRef]
39. Son, N.T.; Chen, C.F.; Cheng, Y.S.; Toscano, P.; Chen, C.R.; Chen, S.L.; Tseng, K.H.; Syu, C.H.; Guo, H.Y.; Zhang, Y.T. Field-scale rice yield prediction from Sentinel-2 monthly image composites using machine learning algorithms. *Ecol. Inform.* **2022**, *69*, 101618. [CrossRef]
40. Prakash, A.; Thangaraj, J.; Roy, S.; Srivastav, S.; Mishra, J.K. Model-Aware XGBoost Method Towards Optimum Performance of Flexible Distributed Raman Amplifier. *IEEE Photonics J.* **2023**, *15*, 8800210. [CrossRef]
41. Lee, H.; Wang, J.; Leblon, B. Using linear regression, random forests, and support vector machine with unmanned aerial vehicle multispectral images to predict canopy nitrogen weight in corn. *Remote Sens.* **2020**, *12*, 2071. [CrossRef]
42. Climate-Data. Climate-Data.org. **2024**. Available online: https://en.climate-data.org/europe/italy/apulia/rutigliano-14028/ (accessed on 7 July 2024).
43. Allen, R.G.; Pereira, L.S.; Raes, D.; Smith, M. *Crop Evapotranspiration-Guidelines for Computing Crop Water Requirements-FAO Irrigation and Drainage Paper 56*; Food and Agriculture Organization of the United Nations: Rome, Italy, **1998**.
44. Tahara, M.; Carver, B.F.; Johnson, R.C.; Smith, E.L. Relationship between relative water content during reproductive development and winter wheat grain yield. *Euphytica* **1990**, *49*, 255–262. [CrossRef]

45. Planet Imagery Product Specifications. **2024**. Available online: www.planet.com/products/satellite-imagery-of-earth/ (accessed on 19 August 2024).
46. Pebesma, E. Simple features for R: Standardized support for spatial vector data. *R J.* **2018**, *10*, 439–446. [CrossRef]
47. Hijmans, R.J.; Bivand, R.; Dyba, K.; Pebesma, E.; Sumner, M.D. CRAN: Package Terra. **2024**. Available online: https://cran.r-project.org/web/packages/terra/index.html (accessed on 11 July 2024).
48. Wu, C.; Niu, Z.; Tang, Q.; Huang, W. Estimating chlorophyll content from hyperspectral vegetation indices: Modeling and validation. *Agric. For. Meteorol.* **2008**, *148*, 1230–1241. [CrossRef]
49. Zhang, Q.; Li, Q.; Zhang, G. Rapid Determination of Leaf Water Content Using VIS/NIR Spectroscopy Analysis with Wavelength Selection. *J. Spectrosc.* **2012**, *27*, 93–105. [CrossRef]
50. Moreira, L.C.J.; Teixeira, A.D.S.; Galvão, L.S. Potential of multispectral and hyperspectral data to detect saline-exposed soils in Brazil. *GIsci. Remote Sens.* **2015**, *52*, 416–436. [CrossRef]
51. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [CrossRef]
52. Gitelson, A.A.; Merzlyak, M.N. Remote sensing of chlorophyll concentration in higher plant leaves. *Adv. Space Res.* **1998**, *22*, 689–692. [CrossRef]
53. Qi, J.; Chehbouni, A.; Huete, A.R.; Kerr, Y.H.; Sorooshian, S. A modified soil adjusted vegetation index. *Remote Sens. Environ.* **1994**, *48*, 119–126. [CrossRef]
54. Rondeaux, G.; Steven, M.; Baret, F. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **1996**, *55*, 95–107. [CrossRef]
55. Sims, D.A.; Gamon, J.A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sens. Environ.* **2002**, *81*, 337–354. [CrossRef]
56. Chen, J.; Jönsson, P.; Tamura, M.; Gu, Z.; Matsushita, B.; Eklundh, L. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sens. Environ.* **2004**, *91*, 332–344. [CrossRef]
57. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [CrossRef]
58. Fitzgerald, G.; Rodriguez, D.; O'Leary, G. Measuring and predicting canopy nitrogen nutrition in wheat using a spectral index—The canopy chlorophyll content index (CCCI). *Field Crops Res.* **2010**, *116*, 318–324. [CrossRef]
59. Kimura, R.; Okada, S.; Miura, H.; Kamichika, M. Relationships among the leaf area index, moisture availability, and spectral reflectance in an upland rice field. *Agric. Water Manag.* **2004**, *69*, 83–100. [CrossRef]
60. Devadas, R.; Lamb, D.W.; Simpfendorfer, S.; Backhouse, D. Evaluating ten spectral vegetation indices for identifying rust infection in individual wheat leaves. *Precis. Agric.* **2009**, *10*, 459–470. [CrossRef]
61. Ju, C.H.; Tian, Y.C.; Yao, X.; Cao, W.X.; Zhu, Y.; Hannaway, D. Estimating Leaf Chlorophyll Content Using Red Edge Parameters. *Pedosphere* **2010**, *20*, 633–644. [CrossRef]
62. Clevers, J.G.P.W.; Gitelson, A.A. Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and -3. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *23*, 344–351. [CrossRef]
63. Verrelst, J.; Schaepman, M.E.; Koetz, B.; Kneubühler, M. Angular sensitivity analysis of vegetation indices derived from CHRIS/PROBA data. *Remote Sens. Environ.* **2008**, *112*, 2341–2353. [CrossRef]
64. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
65. Jiang, J.; Atkinson, P.M.; Chen, C.; Cao, Q.; Tian, Y.; Zhu, Y.; Liu, X.; Cao, W. Combining UAV and Sentinel-2 satellite multi-spectral images to diagnose crop growth and N status in winter wheat at the county scale. *Field Crops Res.* **2023**, *294*, 108860. [CrossRef]
66. Campi, P.; Modugno, A.F.; De Carolis, G.; Pedrero Salcedo, F.; Lorente, B.; Garofalo, S.P. A Machine Learning Approach to Monitor the Physiological and Water Status of an Irrigated Peach Orchard under Semi-Arid Conditions by Using Multispectral Satellite Data. *Water* **2024**, *16*, 2224. [CrossRef]
67. Shah, S.H.; Angel, Y.; Houborg, R.; Ali, S.; McCabe, M.F. A Random Forest Machine Learning Approach for the Retrieval of Leaf Chlorophyll Content in Wheat. *Remote Sens.* **2019**, *11*, 920. [CrossRef]
68. Biau, G.; Scornet, E. A random forest guided tour. *TEST* **2016**, *25*, 197–227. [CrossRef]
69. Joshi, A.; Pradhan, B.; Chakraborty, S.; Behera, M.D. Winter wheat yield prediction in the conterminous United States using solar-induced chlorophyll fluorescence data and XGBoost and random forest algorithm. *Ecol. Inf.* **2023**, *77*, 102194. [CrossRef]
70. Garofalo, S.P.; Modugno, A.F.; De Carolis, G.; Sanitate, N.; Negash Tesemma, M.; Scarascia-Mugnozza, G.; Tekle Tegegne, Y.; Campi, P. Explainable Artificial Intelligence to Predict the Water Status of Cotton (*Gossypium hirsutum* L., 1763) from Sentinel-2 Images in the Mediterranean Area. *Plants* **2024**, *13*, 3325. [CrossRef] [PubMed]
71. Kuhn, M. Journal of Statistical Software Building Predictive Models in R Using the Caret Package. **2008**. Available online: http://www.jstatsoft.org/ (accessed on 7 July 2024).
72. Wright, M.N.; Ziegler, A. Ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *J. Stat. Softw.* **2017**, *77*, 1–17. [CrossRef]

73. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August **2016**; pp. 785–794. [CrossRef]
74. Koc, A.; Odilbekov, F.; Alamrani, M.; Henriksson, T.; Chawade, A. Predicting yellow rust in wheat breeding trials by proximal phenotyping and machine learning. *Plant Methods* **2022**, *18*, 30. [CrossRef] [PubMed]
75. Loi, C.L.; Wu, C.C.; Liang, Y.C. Prediction of Tropical Cyclogenesis Based on Machine Learning Methods and Its SHAP Interpretation. *J. Adv. Model. Earth Syst.* **2024**, *16*, e2023MS003637. [CrossRef]
76. Yuan, X.; Li, S.; Chen, J.; Yu, H.; Yang, T.; Wang, C.; Huang, S.; Chen, H.; Ao, X. Impacts of Global Climate Change on Agricultural Production: A Comprehensive Review. *Agronomy* **2024**, *14*, 1360. [CrossRef]
77. Omia, E.; Bae, H.; Park, E.; Kim, M.S.; Baek, I.; Kabenge, I.; Cho, B.-K. Remote Sensing in Field Crop Monitoring: A Comprehensive Review of Sensor Systems, Data Analyses and Recent Advances. *Remote Sens.* **2023**, *15*, 354. [CrossRef]
78. Peeters, A.; Cohen, Y.; Bahat, I.; Ohana-Levi, N.; Goldshtein, E.; Netzer, Y.; Tenreiro, T.R.; Alchanatis, V.; Ben-Gal, A. A spatial machine-learning model for predicting crop water stress index for precision irrigation of vineyards. *Comput. Electron. Agric.* **2024**, *227*, 109578. [CrossRef]
79. Roy, B.; Sagan, V.; Haireti, A.; Newcomb, M.; Tuberosa, R.; LeBauer, D.; Shakoor, N. Early Detection of Drought Stress in Durum Wheat Using Hyperspectral Imaging and Photosystem Sensing. *Remote Sens.* **2024**, *16*, 155. [CrossRef]
80. Sishodia, R.P.; Ray, R.L.; Singh, S.K. Applications of Remote Sensing in Precision Agriculture: A Review. *Remote Sens.* **2020**, *12*, 3136. [CrossRef]
81. Orlenko, A.; Moore, J.H. A comparison of methods for interpreting random forest models of genetic association in the presence of non-additive interactions. *BioData Min.* **2021**, *14*, 9. [CrossRef] [PubMed]
82. Lin, C.; Doyog, N.D. Challenges of Retrieving LULC Information in Rural Forest Mosaic Landscapes Using Random Forest Technique. *Forests* **2023**, *14*, 816. [CrossRef]
83. Zafar, Z.; Zubair, M.; Zha, Y.; Fahd, S.; Ahmad Nadeem, A. Performance assessment of machine learning algorithms for mapping of land use/land cover using remote sensing data. *Egypt. J. Remote Sens. Space Sci.* **2024**, *27*, 216–226. [CrossRef]
84. Niazkar, M.; Menapace, A.; Brentan, B.; Piraei, R.; Jimenez, D.; Dhawan, P.; Righetti, M. Applications of XGBoost in water resources engineering: A systematic literature review (Dec 2018–May 2023). *Environ. Model. Softw.* **2024**, *174*, 105971. [CrossRef]
85. Devi, K.A.; Rani, K.U. Exploring Machine Learning Algorithms in Soil Management and Precision Agriculture—A Survey. *Int. J. Eng. Res.* **2024**, *13*.
86. Wang, W.; Cheng, Y.; Ren, Y.; Zhang, Z.; Geng, H. Prediction of chlorophyll content in multi-temporal winter wheat based on multispectral and machine learning. *Front. Plant Sci.* **2022**, *13*, 896408. [CrossRef] [PubMed]
87. Zhang, X.; Yang, Z.; Wang, L.; Yue, Y.; Wang, L.; Yang, X. The effects of plant growth-promoting rhizobacteria on plants under temperature stress: A meta-analysis. *Rhizosphere* **2022**, *28*, 100788. [CrossRef]
88. Zhang, X.; Chen, S.; Sun, H.; Pei, D.; Wang, Y. Dry matter, harvest index, grain yield and water use efficiency as affected by water supply in winter wheat. *Irrig. Sci.* **2008**, *27*, 1–10. [CrossRef]
89. Fang, Y.; Wang, J.; Zhang, R.; Li, F.; Liang, L.; Liu, S.; Xu, B.; Chen, Y. Assessing the impact of early and terminal droughts on root growth, grain yield and yield stability in old and modern wheat cultivars on the Loess Plateau. *Agric. Water Manag.* **2024**, *301*, 108940. [CrossRef]
90. Allart, K.; Almoussawi, A.; Kerbey, L.; Catterou, M.; Roger, D.; Mortier, D.; Blanc, E.; Robert, B.; Spicher, F.; Emery, L.; et al. Splitting Nitrogen Fertilization Is More Important than Nitrogen Level When Mixed Wheat Varieties Are Cultivated in a Conservation Agriculture System. *Agronomy* **2023**, *13*, 1295. [CrossRef]
91. Ru, C.; Hu, X.; Wang, W.; Yan, H. Impact of nitrogen on photosynthesis, remobilization, yield, and efficiency in winter wheat under heat and drought stress. *Agric. Water Manag.* **2024**, *302*, 109013. [CrossRef]
92. Xiao, L.; Feng, M.; Yang, W.; Ding, G. Estimation of Water Content in Winter Wheat (*Triticum aestivum* L.) and Soil Based on Remote Sensing Data–Vegetation Index. *Commun. Soil Sci. Plant Anal.* **2015**, *46*, 1827–1839. [CrossRef]
93. Zhu, J.; Lu, J.; Li, W.; Wang, Y.; Jiang, J.; Cheng, T.; Zhu, Y.; Cao, W.; Yao, X. Estimation of canopy water content for wheat through combining radiative transfer model and machine learning. *Field Crops Res.* **2023**, *302*, 109077. [CrossRef]
94. Mokhtari, N.; Majidi, M.M.; Mirlohi, A. Physiological and antioxidant responses of synthetic hexaploid wheat germplasm under drought. *BMC Plant Biol.* **2024**, *24*, 747. [CrossRef]
95. Abdollahpour, S.; Kosari-Moghaddam, A.; Bannayan, M. Prediction of wheat moisture content at harvest time through ANN and SVR modeling techniques. *Inf. Process. Agric.* **2020**, *7*, 500–510. [CrossRef]
96. Torrion, J.A.; Stougaard, R.N. Impacts and Limits of Irrigation Water Management on Wheat Yield and Quality. *Crop Sci.* **2017**, *57*, 3239–3251. [CrossRef]
97. Nowack, J.C.; Atencia-Payares, L.K.; Tarquis, A.M.; Gomez-del-Campo, M. Application of Unmanned Aerial Vehicle (UAV) Sensing for Water Status Estimation in Vineyards under Different Pruning Strategies. *Plants* **2024**, *13*, 1350. [CrossRef] [PubMed]

98. Garofalo, S.P.; Giannico, V.; Costanza, L.; Alhajj Ali, S.; Camposeo, S.; Lopriore, G.; Pedrero Salcedo, F.; Vivaldi, G.A. Prediction of Stem Water Potential in Olive Orchards Using High-Resolution Planet Satellite Images and Machine Learning Techniques. *Agronomy* **2024**, *14*, 1. [CrossRef]
99. Colorado, J.D.; Calderon, F.; Mendez, D.; Petro, E.; Rojas, J.P.; Correa, E.S.; Mondragon, I.F.; Rebolledo, M.C.; Jaramillo-Botero, A. A novel NIR-image segmentation method for the precise estimation of above-ground biomass in rice crops. *PLoS ONE* **2020**, *15*, e0239591. [CrossRef] [PubMed]
100. Elliott, G.A.; Regan, K.L. Use of reflectance measurements to estimate early cereal biomass production on sandplain soils. *Aust. J. Exp. Agric.* **1993**, *33*, 179–183. [CrossRef]
101. Li, J.; Shi, Y.; Veeranampalayam-Sivakumar, A.N.; Schachtman, D.P. Elucidating sorghum biomass, nitrogen and chlorophyll contents with spectral and morphological traits derived from unmanned aircraft system. *Front. Plant Sci.* **2018**, *9*, 1406. [CrossRef] [PubMed]
102. Haboudane, D.; Miller, J.R.; Tremblay, N.; Zarco-Tejada, P.J.; Dextraze, L. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* **2002**, *81*, 416–426. [CrossRef]
103. Braga, P.; Crusiol, L.G.T.; Nanni, M.R.; Caranhato, A.L.H.; Fuhrmann, M.B.; Nepomuceno, A.L.; Neumaier, N.; Farias, J.R.B.; Koltun, A.; Gonçalves, L.S.A.; et al. Vegetation indices and NIR-SWIR spectral bands as a phenotyping tool for water status determination in soybean. *Precis. Agric.* **2021**, *22*, 249–266. [CrossRef]
104. Zahir SA, D.M.; Omar, A.F.; Jamlos, M.F.; Azmi, M.A.M.; Muncan, J. A review of visible and near-infrared (Vis-NIR) spectroscopy application in plant stress detection. *Sens. Actuators A Phys.* **2021**, *338*, 113468. [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.