



## Article

# Predicting Grapevine Physiological Parameters Using Hyperspectral Remote Sensing Integrated with Hybrid Convolutional Neural Network and Ensemble Stacked Regression

Prakriti Sharma <sup>1</sup>, Roberto Villegas-Díaz <sup>2</sup> and Anne Fennell <sup>1,\*</sup>

<sup>1</sup> Agronomy, Horticulture and Plant Science Department, South Dakota State University, Brookings, SD 57007, USA; prakriti.sharma@sdsu.edu

<sup>2</sup> Department of Public Health, Policy and Systems, University of Liverpool, Liverpool L69 3GL, UK; r.villegas-diaz@liverpool.ac.uk

\* Correspondence: anne.fennell@sdsu.edu

**Abstract:** Grapevine rootstocks are gaining importance in viticulture as a strategy to combat abiotic challenges, as well as enhance scion physiology. Direct leaf-level physiological parameters like net assimilation rate, stomatal conductance to water vapor, quantum yield of PSII, and transpiration can illuminate the rootstock effect on scion physiology. However, these measures are time-consuming and limited to leaf-level analysis. This study used different rootstocks to investigate the potential application of aerial hyperspectral imagery in the estimation of canopy level measurements. A statistical framework was developed as an ensemble stacked regression (REGST) that aggregated five different individual machine learning algorithms: Least absolute shrinkage and selection operator (Lasso), Partial least squares regression (PLSR), Ridge regression (RR), Elastic net (ENET), and Principal component regression (PCR) to optimize high-throughput assessment of vine physiology. In addition, a Convolutional Neural Network (CNN) algorithm was integrated into an existing REGST, forming a hybrid CNN-REGST model with the aim of capturing patterns from the hyperspectral signal. Based on the findings, the performance of individual base models exhibited variable prediction accuracies. In most cases, Ridge Regression (RR) demonstrated the lowest test Root Mean Squared Error (RMSE). The ensemble stacked regression model (REGST) outperformed the individual machine learning algorithms with an increase in  $R^2$  by (0.03 to 0.1). The performances of CNN-REGST and REGST were similar in estimating the four different traits. Overall, these models were able to explain approximately 55–67% of the variation in the actual ground-truth data. This study suggests that hyperspectral features integrated with powerful AI approaches show great potential in tracing functional traits in grapevines.

**Keywords:** aerial hyperspectral remote sensing; deep learning; ensemble learning; stacking regressor; UAV; high throughput; phenotyping; grapevine; rootstock; carbon assimilation; stomatal conductance; transpiration



**Citation:** Sharma, P.; Villegas-Díaz, R.; Fennell, A. Predicting Grapevine Physiological Parameters Using Hyperspectral Remote Sensing Integrated with Hybrid Convolutional Neural Network and Ensemble Stacked Regression. *Remote Sens.* **2024**, *16*, 2626. <https://doi.org/10.3390/rs16142626>

Academic Editor: András Jung

Received: 8 May 2024

Revised: 4 July 2024

Accepted: 9 July 2024

Published: 18 July 2024



**Copyright:** © 2024 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

Recent climate change studies show significant impacts on viticulture production due to extreme climate phenomena, rising temperatures, drought, and increased atmospheric CO<sub>2</sub> concentrations [1,2]. Several studies show that higher temperatures result in the modification of grapevine phenological stages, potentially altering the timing of budbreak, bloom, and fruit ripening [3,4]. Examination of historical data (1981–2007) for temperature influence on the phenological stages suggests that climate change accounts for about 26% of the variability in phenological timing and an earlier transition of the fruit ripening period into warmer conditions [5]. Similarly, the increasing CO<sub>2</sub> levels are associated with

a decrease in stomatal conductance and, consequently, transpiration, which may lead to an increase in biomass in grapevines under conditions of reduced water loss [6]. An increase in temperature and CO<sub>2</sub> is also correlated with a decline in photosynthesis due to nitrogen depletion [7].

One long-term strategy to address these issues involves using climate-resilient scion-rootstock combinations to gain a deeper understanding of their physiological dynamics. Grafting onto rootstocks in grapevines was initially used to overcome the root-damaging insect pest *Phylloxera* (*Daktulosphaira vitifoliae*) in the mid-1800s [8–10]. Nevertheless, with changing climatic conditions, the selection of appropriate rootstock genotype has gained a new impetus for improving scion physiology, mobilizing resource-use efficiency, and adapting grapevine to diverse biotic, and abiotic stressors [11,12]. It is known that traits related to berry production, such as conferred scion vigor, water-use efficiency, net assimilation of carbon, leaf stomatal conductance, and hydraulic root–shoot signaling, are influenced by rootstock [13–25].

At present, in economically important perennials, rootstocks have been used to combat soil-borne biotic and abiotic stressors; however, less is known about their impact on scion physiology and molecular interactions [12]. Likewise, rootstock selection through rapid, repeatable, and large-scale phenotyping for physiological traits is challenging. Traditional phenotyping procedures are low-throughput and involve destructive sampling. Such phenotyping approaches for physiological parameters rely on manual, laborious, time-intensive, and inefficient equipment, leading to subjective and inaccurate data. Therefore, effective monitoring systems are needed to accurately measure key variables at high spatial and temporal resolution for a reliable understanding of vine physiology.

Recently, hyperspectral imagery has been used to monitor physiological traits like CO<sub>2</sub> assimilation [26], photosynthesis [27,28], and plant water status and transpiration [29,30] in grapevines. In comparison to multispectral sensors, a hyperspectral sensor can measure the radiative properties of a plant through numerous contiguous narrow wavelength bands. These narrow wavelength bands can detect subtle variations in the physiology of plants as compared to broad bands [31,32]. However, hyperspectral sensor-derived narrow bands are correlated and may have repetitive information, resulting in substantial noise [31]. In addition, developing predictive models using such spectral features is difficult as the amount of directly measured physiological (ground-truth) data is limited.

Several machine learning (ML) modeling approaches are applied to correlate hyperspectral reflectance measures to ground-truth physiological measurements [33,34]. The widely used model algorithms for hyperspectral data are based on dimensionality reduction of spectra. Partial least square regression (PLSR) is a modified form of linear regression, which can provide robust prediction in cases where the number of predictors is greater than the number of samples [35,36]. In particular, when predictor variables have a high degree of correlation, this technique works very well at addressing multicollinearity problems [37]. The *M* linear combinations, or “PLS components”, of the original *p* predictor variables are computed as part of the PLSR working principle. These *M* linear combinations are then used as predictors in a linear regression model and fitted using the least squares approach. Similarly, the principal component regression (PCR) algorithm is another dimension reduction technique through principal component analysis, which results in a smaller set of uncorrelated predictor variables [38]. Subsequently, linear regression is performed with the principal components as predictors or features instead of the original variables. This approach potentially addresses multicollinearity as well as dimensionality constraints, hence improving model predictive performance.

Other regularization-based algorithms like least absolute shrinkage and selection operator (Lasso), ridge regression (RR), and elastic net (ENET), which address multicollinearity and handling of high-dimension data formats [39,40], are also widely used for hyperspectral data analysis. The Lasso algorithm effectively adds a penalty component to the linear regression cost function. This penalty term is based on a regularization technique in which the regularization parameter  $\lambda$  can result in coefficient shrinkage. This

is critical in high-dimensional datasets since the technique reduces the coefficient of less important features to zero. As a result, it performs feature selection and can address multicollinearity difficulties by eliminating the correlated variables [41]. Like Lasso, the ridge regression (RR) algorithm is based on regularization techniques in linear regression. The main difference is that RR adds squared values of the coefficients of predictor variables of the cost function instead of absolute values used in Lasso. In contrast, the regularization parameter  $\lambda$  does not shrink the coefficient to zero in RR unlike Lasso; instead, it tends to make the coefficient of less important variables minimal [42]. Therefore, there is no feature selection in RR, and it downplays the impact of less important features. The elastic net (ENET) algorithm combines aspects of both Lasso and RR. More specifically, it adds and balances a linear combination of L1 (Lasso) and L2 (RR) penalties to the cost function. Therefore, ENET often results in features selected where coefficients are shrunk but not eliminated. This is particularly useful when dealing with high-dimensional features, as it retains a balance between feature selection and coefficient shrinkage [43].

Performance varies greatly, with the application of different models for predicting physiological traits, depending on the principle of the algorithm being used. Therefore, it is reasonable to aggregate different performances using a heterogeneous ensemble approach like ensemble stacking, which integrates the strengths of multiple machine learning models. Fu et al. (2019) [44] combines six different ML techniques as base learners and predicts photosynthetic parameters ( $J_{\max}$  and  $V_{\max}$ ) using hyperspectral leaf data. The ensemble stacked regression (REGST) algorithm is a powerful machine learning technique that combines the prediction of multiple individual models to generate an overall precise model. It involves multiple base models that are hyper-tuned and trained in the same dataset. The predictions derived from these base models are the input feature for a meta-model, which essentially learns to weigh the prediction of each base model to generate final ensemble predictions [45].

Hybrids of these individual models can also be used; however, despite many studies with single or ensemble-based ML techniques, very few findings have been published regarding hybrid algorithms where deep learning (DL) is utilized as a feature extractor in combination with other ML techniques. This hybrid approach allows gathering the strength of DL for capturing relevant information from hyperspectral signals and passes these learned features as input into ML algorithms in the case of limited data scenarios. Most of the publications related to hybrid models are limited to varietal or biotic/abiotic stress classification cases, where a convolutional neural network (CNN) has been used as a feature extractor combined with ML techniques, which have shown improved accuracy over ML models alone [46,47]. The CNN algorithm is specifically designed to handle a one-dimensional sequence of data. Applying convolutions in a single dimension captures the pattern or relationship in data, which is very powerful for tasks like hyperspectral signal processing. The different components of CNN are convolutional layers, pooling layers, and fully connected layers [48]. The convolutional layers have a filter component that convolves around the input sequence to extract features. The use of pooling layers reduces the number of features and spatial dimensionality before feeding to the next layer. Fully connected or dense layers are generally used to make final predictions based on learned features. Within the field of physiological and AI-integrated studies, this research offers a unique approach by combining the REGST model with modern DL techniques like CNN. These strategies work together to automate the feature extraction process from hyperspectral data, distinguishing it from traditional approaches. This type of hybrid model is often used in limited data scenarios since CNN alone requires high-sample data for robust prediction [49].

The research employs statistical analyses to compare ground-truth physiological measurements across various scion/rootstock combinations as foundational information. The study then explores the potential application of hyperspectral data and artificial intelligence for quantifying physiological traits. Specifically, it addresses the following questions: (1) Are hyperspectral data capable of providing an indirect measure of grapevine physiological parameters? (2) Does an ensemble stacked regression model work better in learning and predicting physiological traits from hyperspectral data than individual ML models? (3) Does adding a DL algorithm to extract hyperspectral features to an existing REGST improve model prediction accuracy? To answer these queries the following approach was developed: (1) Aerial-based hyperspectral data were retrieved for individual grafted vines; and (2) the potential for prediction of ground-truth photosynthetic parameters was explored using five different ML algorithms (PLSR, Lasso, ENET, RR, CNN), ensemble stacked regression (REGST), and a proposed hybrid model—the combination of convolutional neural network and ensemble stacked regression (CNN-REGST).

## 2. Materials and Methods

### 2.1. Site and Experimental Design

Vine physiological and hyperspectral profiles were measured in the field at the South Dakota State University research vineyard in Brookings, SD (44.311356, −96.798386). *Vitis* hybrid ‘Marquette’ grafted to five commercial rootstocks 1103 Paulsen (1103P), 3309 Couderc (3309 C), Teleki 5C (5C), Freedom, Selection Oppenheim 4 (SO4), and a ‘Marquette’ (homograft as control) were used to measure rootstock influence on scion physiology. A block of these scion/rootstock combinations was placed in each row resulting in four replicates per rootstock (Supplementary Table S1). Five-year-old vines, in east–west row orientation, with a spacing of 1.828 m and 3.048 m between vines and rows, respectively. Vines were trained in a high-cordon management system. All measurements were taken in 2021 and 2022 growing seasons. The monthly mean temperature throughout the growing season (May–October) was 18.51 °C and 18.23 °C in 2021 and 2022, respectively. Average monthly humidity for the growing season was 71% and 70.6% for 2021 and 2022, respectively.

### 2.2. Directly Measured Physiological Attributes

Given that the rootstock can potentially influence water-use efficiency and photosynthetic traits, the following variables were targeted as ground-truth data: net assimilation rate (A), stomatal conductance to water vapor (gsw), transpiration rate (E), and effective quantum yield of PSII in light adaptive stage ( $\phi$  PSII). These data were acquired using a portable photosynthesis system LI-6800 (LI-COR Biosciences, Lincoln, NE, USA) (Figure 1A). The LI-COR settings were fixed for the temporal measurements: Flow rate of 600  $\mu\text{mol s}^{-1}$ , temperature and relative humidity set closest to ambient conditions, reference  $\text{CO}_2$  to 400  $\mu\text{mol mol}^{-1}$ , and saturating light of 1800  $\mu\text{mol m}^{-2} \text{s}^{-1}$ . Measurements were conducted three separate times in 2021 (12 June, 21 July, and 14 August) and four separate times in 2022 (21 July, 17 August, 4 September, and 17 October) during daylight hours from 9 am to 12 pm. A single leaf from each vine was selected for measurement, resulting in four replicate samples per genotype (6 genotypes), resulting in a total of 24 samples for each sampling date. The selected leaves were fully developed, healthy middle leaves that had adapted to sunlight conditions ensuring consistency in the physiological measurements and minimizing potential variability due to leaf development stages or environmental factors. Analysis of the directly measured traits was conducted by combining the two years. First, these directly measured traits were transformed to meet normality assumptions and then a one-way ANOVA analysis was conducted to compare means across genotypes.





**Figure 1.** (A) Fully developed, sun-exposed leaves measured by LiCOR 6800 as ground-truth or directly measured variables. (B) UAV platform (DJI Matrice 600 drone) for indirect measurement of physiological traits through aerial remote sensing. (C) Nano hyperspectral headwall sensor fixated in gimbal carried by UAV.

### 2.3. Hyperspectral Data Acquisition and Preprocessing

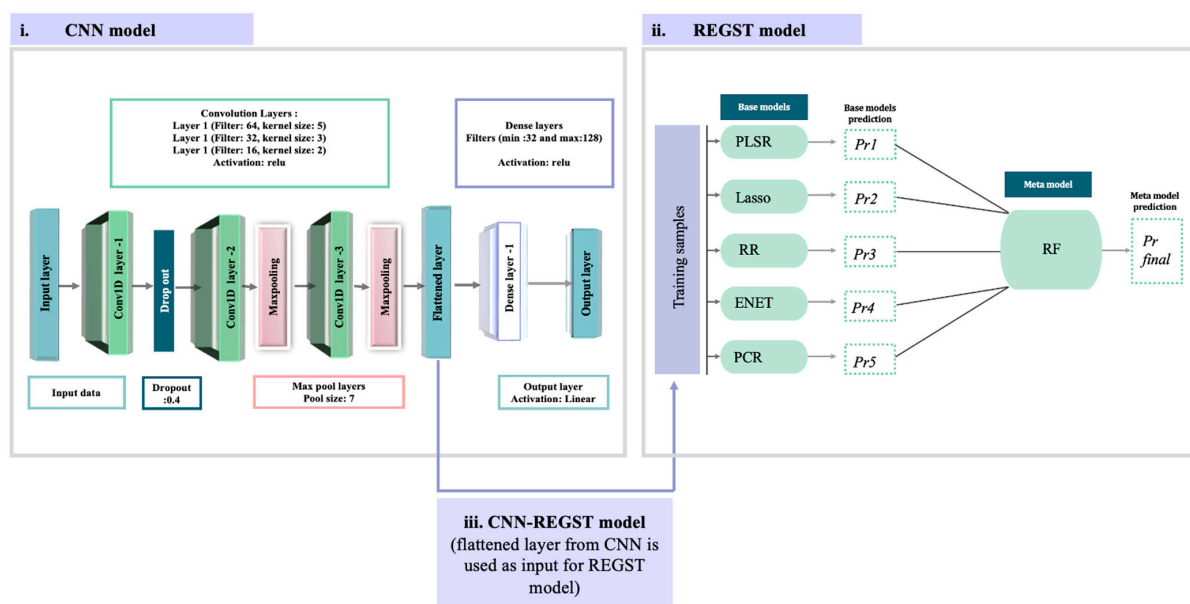
The vine spectral characteristics were obtained using the Headwall Nano-Hyperspec® push-broom hyperspectral imager (Figure 1B,C), a VNIR (visible–near-infrared) sensor (Headwall Photonics Inc., Bolton, MA, USA). The sensor was equipped with an 8 mm lens, offering a  $30.4^\circ$  field of view (FOV). It gathered data in 271 bands with a 2.2 nm pixel spacing and 12-bit radiometric resolution, featuring a full width half maximum (FWHM) of 6 nm. To ensure data quality and minimize external disturbances like roll, pitch, and yaw oscillations, the sensor was integrated into a custom gimbal system. This complete setup was integrated into the DJI Matrice 600 Pro six-rotor UAV (Shenzhen Dajiang Innovation Technology Co., Ltd., Shenzhen, China). The UAV flight was conducted at a consistent altitude in both years, resulting in an image with a spatial resolution of 1.87 cm per pixel. Also, flight time and schedule were arranged consecutively with the LI-COR direct sampling.

The radiometric calibration and geometric correction processing were performed using the HyperSpec III (v3.1.5.1) and SpectralView (v3.1.5.1) software from Headwall Photonics Inc. (Fitchburg, MA, USA). The software utilized sensor-specific response information to process radiance and a dark reference image. Subsequently, the reflectance values were computed using the empirical line conversion method [50] with the aid of a reflectance standard. The orthomosaic images were then processed using ENVI (version 5.7) software (L3Harris Geospatial Solutions Inc., Broomfield, CO, USA). The canopy of each vine was saved in shapefile format, which was used to derive spectral signatures. To derive the spectral signature, only the vine vegetation surface was considered to create shapefile polygons. This shapefile encompassed the canopy area of the target grapevine, defined as a region of interest (ROI), which varied depending on the canopy size. Average pixel values within this ROI were then used for further analysis. Spectral features were extracted using ‘Rasterio’ library in Python [51]. Hyperspectral data spectral correction was performed to address noise or remove spikes and drops. Subsequently, they were processed through the Savitzky–Golay filter using the Python ‘SciPy’ module with a window width of 7 and second-order polynomial smoothing [52,53].

## 2.4. Modeling Assessment

### 2.4.1. Model Algorithm Selection

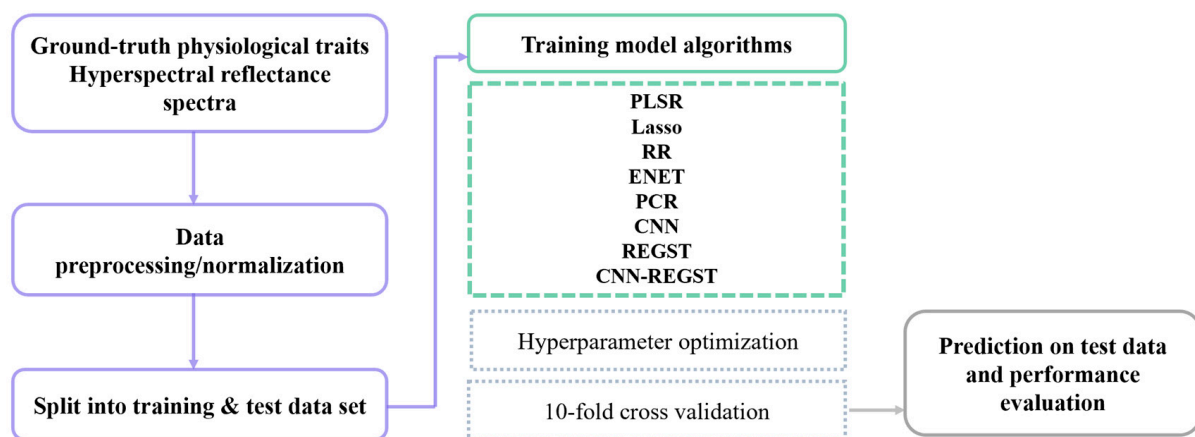
Gathering sufficient ground-truth physiological measurements is challenging due to measurement time constraints, resulting in a limited number of samples. As a result, the data samples ( $n$ ) collected are limited compared to the high-dimensional spectral features predictor variables ( $p$ ). This leads to overfitting, poor generalizations, and other issues in modeling. Hence, modeling approaches in this study were entirely selected based to address this constraint. The models used were individual models: PLSR [37]; Lasso [41]; RR [42]; ENET [43]; PCR [38]; CNN model [48]; a proposed ensemble stacked regression model (REGST), and a hybrid model (CNN-REGST), which combined the CNN and REGST. The individual models, PLSR, Lasso, RR, ENET, and PCR, were selected based on either projection-based principles (PLSR and PCR) or regularization-based principles (Lasso, RR, and ENET) with the aim of dimensionality reduction. CNN as the individual model was designed with three convolutional layers and one dense layer as illustrated in Figure 2(i). REGST model is illustrated in Figure 2(ii), where the base model includes PLSR, Lasso, RR, ENET, and PCR. In the first stage, the base models were fine-tuned and fed to the meta-model in the second stage to make a final prediction (REGST). REGST was evaluated for each physiological parameter separately. The combination of CNN and REGST model was proposed for this study to leverage the strength of CNN as a feature extractor (Figure 2(iii)) and use that information as a predictor for REGST model. A similar architecture to CNN as individual model was applied in the context of layers, filter size, and other parameters, except the flattened layer was replaced with REGST model to form the CNN-REGST model (Figure 2). The CNN-REGST model was evaluated for each physiological parameter separately.



**Figure 2.** Workflow designed for (i) 1D CNN model architecture. The model consists of 3 different convolutional layers and two max-pooling layers that are flattened and forwarded to a dense layer to estimate the final output. (ii) Ensemble stacked regression model (REGST), where partial least squares regression (PLSR), Least Absolute Shrinkage and Selection Operator (LASSO), ridge regression (RR), elastic net (ENET), and principal component regression (PCR) are used as base models, and random forest regression (RF) is used as a meta-model to make final predictions on different physiological parameters. (iii) Proposed hybrid CNN-REGST; the model was designed using CNN for feature extraction and REGST for the regression task. The initial CNN operates on one-dimensional hyperspectral data, and the flattened layer obtained is forwarded as input for the REGST model to make predictions for different physiological traits.

#### 2.4.2. Modeling Pipeline and Hyperparameter Optimization

The model algorithms were analyzed in three different aspects: (1) analyzing all base models and CNN as individual separate models; (2) REGST model analysis; and (3) CNN-REGST model analysis. As illustrated in Figure 3, all the models imported the directly measured (ground-truth) physiological traits as a response and transformed hyperspectral reflectance data as predictor variables. The random split of the whole two-year dataset was done into two parts: training (70%) and testing (30%). All the individual base models were trained using the ‘scikit-learn’ package in Python 3.6 [54]. The different hyperparameters listed in Table 1 were tuned using grid search with 10-fold cross-validation [55]. The grid parameters used for each hyperparameter are listed in Supplementary Table S2. Then, the final model with the optimal hyperparameter set was used in test dataset, and the prediction performance of each individual model was analyzed using various model metrics. For CNN implementation (Figure 2(i)) RandomSearch based on ‘keras’ package in Python 3.6 [56] was used to select the learning rate for modeling each physiological parameter. Likewise, the number of epochs and batch size were set based on the lowest MSE obtained in the training phase with 0.1 validation split. In the case of the REGST model, (Figure 2(ii)) StackingRegressor function was utilized from ‘scikit-learn’ library [54]. The hyperparameters for the base models used in REGST were decided based on optimal hyperparameters obtained during individual base model assessment. For RF metamodel, sets of optimal hyperparameters were achieved using RandomSearch with 10-fold cross-validation [57].



**Figure 3.** Model pipeline with cross-validation approach for predicting net assimilation rate, stomatal conductance to water vapor, quantum yield of PSII, and transpiration.

**Table 1.** Hyperparameters used for the model algorithms.

Model Name	Model Hyperparameters
PLSR	n_components
Lasso	alpha, fit_intercept
RR	alpha, fit_intercept
ENET	alpha, fit_intercept, l1_ratio
PCR	n_components
RF	max_depth, min_samples_leaf, min_samples_split, n_estimators
CNN	Learning rate, batch_size, epochs, activation function

### 2.4.3. Evaluation Metrics

For comparing model performances, three evaluation metrics were derived for both training and test data prediction, root mean square error (RMSE), mean absolute error (MAE), and coefficient of determination ( $R^2$ ). The RMSE refers to the squared root of the squared difference in observed and predicted values. The MAE is the measure of the absolute difference between predicted and actual values. The coefficient of determination is the fraction of variation in response variables that can be predicted by dependent variables. It is calculated by dividing the sum of squares residuals by the total sum of squares and the value varies from 0 to 1.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

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

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

### 2.4.4. Permutation-Based Feature Importance Score

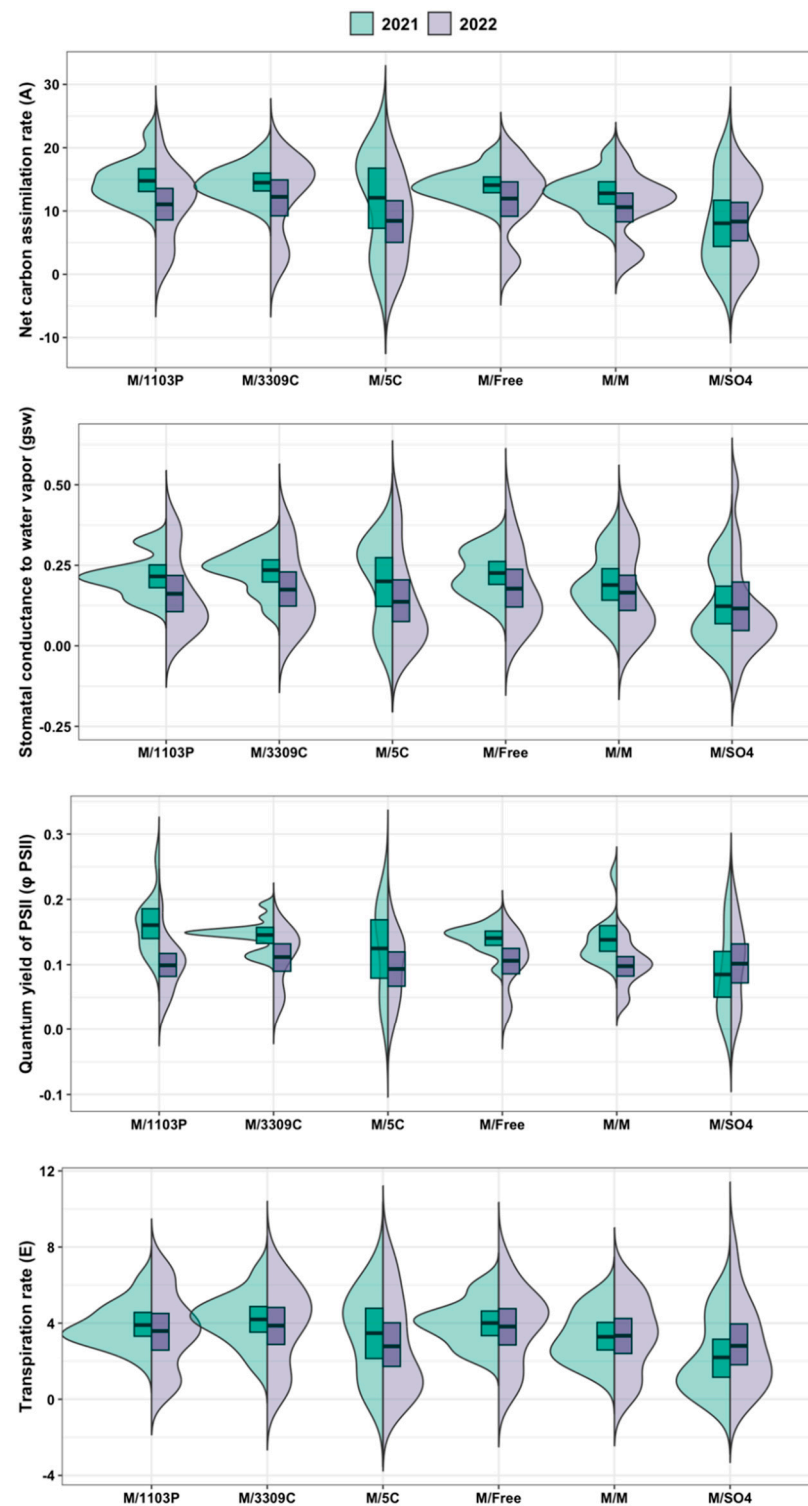
Permutation-based feature performance is an important measure that helps assess the role of each feature in model performance. This is calculated by observing test model performance when the values of specific features are randomly shuffled [58]. While shuffling specific features and keeping other data intact, it captures interactions amongst the features. This is a model-agnostic approach that can be applied to any ML algorithm like linear or tree-based models. This was applied using ‘mlxtend’ library [59] in Python using feature importance permutation.

## 3. Results

### 3.1. Directly Measured Leaf Mesophyll Traits and Physiology

The distribution of different ground-truth leaf measurement values is shown in the violin plot in Figure 4. The 2021 data were normally distributed in contrast to 2022. One reason for this is an additional sampling occurred in 2022 only, near the end of the growing season, and all those measurements were lower in value than the earlier measures. Comparing leaf mesophyll and physiological traits among the different genotypes, M/5C showed a wide range of values (Table 2). The net assimilation rate, stomatal conductance to water vapor, and transpiration rate of ‘Marquette’ grafted with rootstocks 1103P, 3309C, and Freedom were found to be higher than those of the 5C and SO4 genotypes, as indicated in Table 2. The performance of the homograft (M/M) exhibited a moderate level of physiological efficacy, falling between rootstock genotypes that had higher and lower performance. The results from ANOVA analysis showed that genotypes were significantly different in their net assimilation rate, stomatal conductance to water vapor, and transpiration rate. A post hoc Tukey test showed the influence of rootstock genotypes for each physiological parameter. The results indicated that there was no significant difference in net assimilation rate and stomatal conductance to water vapor between grafts with rootstock 3309C, 1103P, and Freedom, whereas SO4 differed statistically from all three of these rootstocks. In the case of transpiration rate, M/3309C had the highest average value, whereas M/SO4 had the lowest.





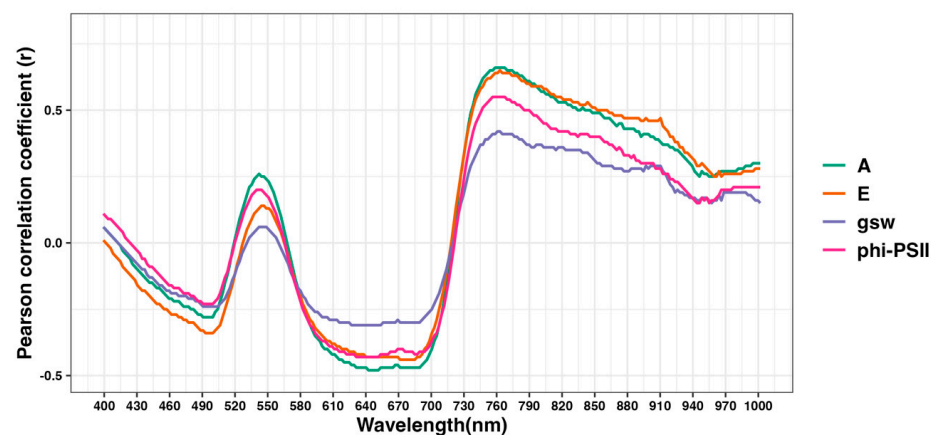
**Figure 4.** Distribution of physiological measures in scion as conferred by six different rootstocks. A = net assimilation rate,  $\mu\text{mol m}^{-2} \text{s}^{-1}$ ; gsw = stomatal conductance to water vapor,  $\text{mol m}^{-2} \text{s}^{-1}$ ; PSII, effective quantum yield of PSII in light adaptive stage, E = transpiration rate,  $\text{mmol m}^{-2} \text{s}^{-1}$  recorded in 2021 and 2022.

**Table 2.** Descriptive statistics of the ground-truth physiological measures for ‘Marquette’ grafted to six different rootstocks. A = net assimilation rate,  $\mu\text{mol m}^{-2} \text{s}^{-1}$ ; gsw = stomatal conductance to water vapor,  $\text{mol m}^{-2} \text{s}^{-1}$ ; PSII, effective quantum yield of PSII in light adaptive stage, E = transpiration rate,  $\text{mmol m}^{-2} \text{s}^{-1}$ . The data presented in the table are the average trait values measured over two years with standard error. Significant trait differences were identified with ANOVA ( $p$ -value  $< 0.05$ ), and differences between rootstock mean values are identified by different letters within a given trait.

Genotypes	A	gsw	$\phi$ PSII	E
M/1103P	12.64 a $\pm$ 0.95	0.18a $\pm$ 0.01	0.12 $\pm$ 0.009	3.72ab $\pm$ 0.31
M/3309C	13.19 a $\pm$ 0.92	0.20a $\pm$ 0.01	0.12 $\pm$ 0.007	4.01a $\pm$ 0.33
M/5C	09.83 ab $\pm$ 1.45	0.16ab $\pm$ 0.02	0.10 $\pm$ 0.012	3.04ab $\pm$ 0.48
M/Freedom	12.87 a $\pm$ 0.88	0.19a $\pm$ 0.01	0.12 $\pm$ 0.007	3.90ab $\pm$ 0.31
M/M	11.54 ab $\pm$ 0.82	0.17a $\pm$ 0.02	0.11 $\pm$ 0.007	3.31ab $\pm$ 0.31
M/SO4	08.21 b $\pm$ 1.22	0.11b $\pm$ 0.02	0.09 $\pm$ 0.012	2.53b $\pm$ 0.40

### 3.2. Relationship of Directly Measured Vine Traits with Hyperspectral Features

A Pearson correlation was determined between directly measured plant traits and each reflectance wavelength (400–1000 nm). The line plots (Figure 5) illustrate the Pearson correlation coefficient ‘r’ value for respective traits. Both positive and negative correlations were observed between hyperspectral data and directly measured plant traits. The pattern of each curve resembles the reflectance signature of the correlation between the vine’s direct physiological trait measure and reflectance (indirect measure). The first peak covers the wavelength range from 520 to 565 nm, which is the visible green spectrum, and it has a positive correlation with directly measured traits. Likewise, a second peak having wavelength range of 730–790 nm has a positive correlation ( $r \gg 0.5$ ) with different measured physiological traits except for stomatal conductance to water vapor (gsw). While some negative Pearson correlation coefficient values in form of lowest points were also observed at around 580–700 nm (visible yellow to red).



**Figure 5.** Relationship between hyperspectral data and various physiological traits. Correlation coefficient (r) for net carbon assimilation (A, green)  $\mu\text{mol m}^{-2} \text{s}^{-1}$ ; transpiration (E, orange)  $\text{mmol m}^{-2} \text{s}^{-1}$ ; stomatal conductance to water vapor (gsw, blue)  $\text{mol m}^{-2} \text{s}^{-1}$ ; and quantum yield of PSII ( $\phi$  PSII, magenta).

### 3.3. Performance of Different Modeling Algorithms

Physiological trait prediction performance of each regression algorithm is shown in Table 3(i–iv). These are the combined evaluation metrics for the training and test dataset. For every trait, the best prediction model was selected based on lowest RMSE. The test performance of each model for predicting different physiological trait values are shown in Supplementary Figures S1–S4.

**Table 3.** Model algorithm performance of base models (PLSR, Lasso, RR, ENET, PCR), ensemble stacked regression model (REGST), hybrid model combining CNN, and ensemble stacked regression (CNN-REGST) model used for hyperspectral prediction of i. Net assimilation rate, ii. stomatal conductance to water vapor, iii. quantum yield of PSII, and iv. transpiration rate across two growing seasons. Performance measures are mean absolute error (MAE), root mean squared (RMSE), and coefficient of determination ( $R^2$ ).

i. Performance results for net assimilation rate (A)						
Model name	Training performance (n = 114)			Test Performance (n = 49)		
	MAE	RMSE	$R^2$	MAE	RMSE	$R^2$
PLSR	2.18	2.84	0.72	2.74	3.95	0.58
Lasso	2.24	2.93	0.70	2.88	4.16	0.53
RR	2.07	2.69	0.74	2.76	3.81	0.61
ENET	2.22	2.92	0.70	2.89	4.14	0.53
PCR	2.18	2.87	0.71	2.81	3.90	0.59
REGST	2.01	2.51	0.81	2.70	3.51	0.64
CNN	3.04	3.89	0.61	3.07	3.89	0.51
CNN-REGST	2.41	2.95	0.71	2.68	3.55	0.65
ii. Performance results for stomatal conductance to water vapor (gsw)						
Model name	Training performance (n = 114)			Test Performance (n = 49)		
	MAE	RMSE	$R^2$	MAE	RMSE	$R^2$
PLSR	0.04	0.05	0.72	0.06	0.08	0.57
Lasso	0.04	0.05	0.71	0.06	0.08	0.55
RR	0.03	0.05	0.77	0.07	0.09	0.55
ENET	0.04	0.05	0.71	0.06	0.09	0.55
PCR	0.05	0.06	0.62	0.07	0.09	0.49
REGST	0.04	0.05	0.74	0.06	0.08	0.58
CNN	0.06	0.09	0.42	0.06	0.08	0.37
CNN-REGST	0.05	0.06	0.62	0.06	0.08	0.61
iii. Performance results for quantum yield of PSII ( $\phi$ PSII)						
Model name	Training performance (n = 114)			Test Performance (n = 49)		
	MAE	RMSE	$R^2$	MAE	RMSE	$R^2$
PLSR	0.02	0.02	0.68	0.03	0.04	0.46
Lasso	0.02	0.02	0.68	0.02	0.04	0.52
RR	0.02	0.02	0.67	0.03	0.04	0.51
ENET	0.02	0.03	0.60	0.03	0.04	0.48
PCR	0.02	0.02	0.67	0.03	0.04	0.46
REGST	0.02	0.02	0.69	0.02	0.04	0.54
CNN	0.02	0.03	0.39	0.03	0.04	0.33
CNN-REGST	0.02	0.02	0.65	0.02	0.04	0.55
iv. Performance results for transpiration rate (E)						
Model name	Training performance (n = 114)			Test Performance (n = 49)		
	MAE	RMSE	$R^2$	MAE	RMSE	$R^2$
PLSR	0.81	0.99	0.70	1.13	1.39	0.55
Lasso	0.75	0.92	0.74	1.07	1.32	0.59
RR	0.75	0.91	0.75	0.96	1.21	0.65
ENET	0.75	0.92	0.74	0.98	1.23	0.64
PCR	0.85	1.06	0.66	1.13	1.41	0.53
REGST	0.68	0.87	0.77	0.91	1.19	0.67
CNN	0.93	1.17	0.60	1.35	1.63	0.38
CNN-REGST	0.80	0.98	0.71	0.98	1.19	0.67

For net assimilation rate, the training  $R^2$  varied between 0.61 and 0.81, while the RMSE error ranged from 2.51 to 3.89 (Table 3(i)). Among different training models, REGST outperformed the rest with the lowest RMSE. In contrast to training, the results obtained in testing phase had higher RMSE value with lower  $R^2$ . Among the various individual base models evaluated in testing phase, the RR model had lowest RMSE. Comparatively, the REGST and hybrid CNN-REGST models performed well with the lowest RMSE.

For stomatal conductance to water vapor, the RMSE and  $R^2$  values in the training phase ranged from 0.5 to 0.9 and 0.42 to 0.77, respectively (Table 3(ii)). The training performance

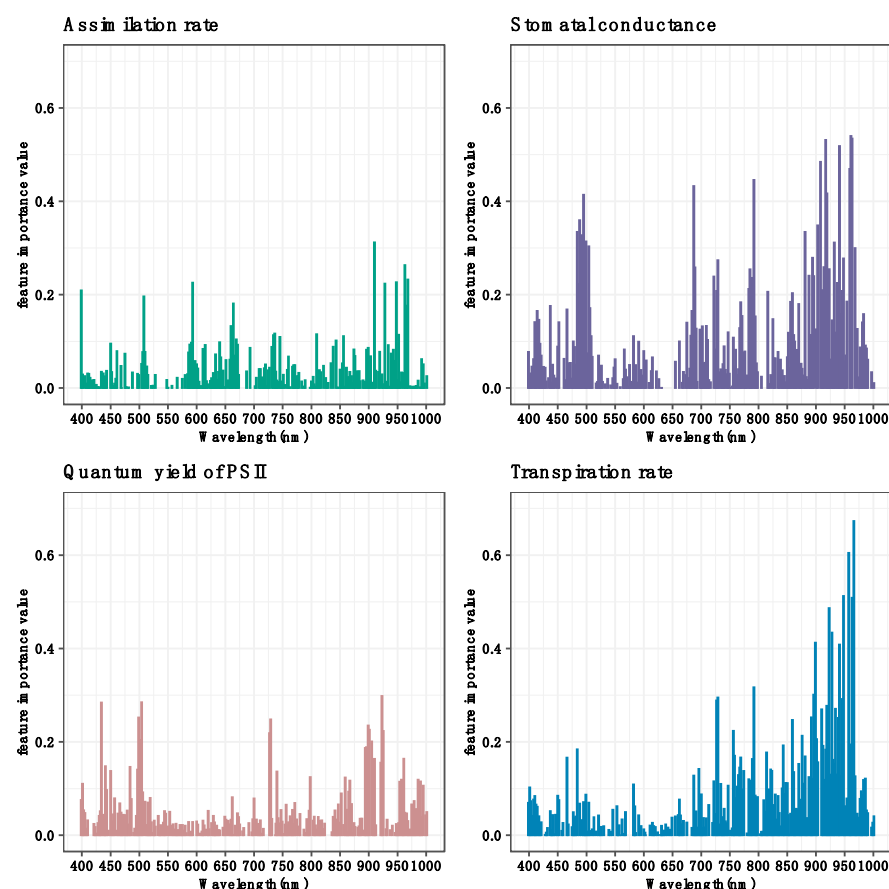
of all individual base models, as well as REGST and CNN-REGST, had the lowest error in contrast to the CNN model. In the testing phase, compared to individual base models, the REGST model performed better with a lower RMSE and greater  $R^2$ . CNN-REGST exhibited a similar performance to that of REGST as indicated by the RMSE and  $R^2$ .

For the quantum yield of PSII, nearly all models exhibited comparable prediction accuracies during both the training and test phases, except for the CNN model (Table 3(iii)). In the training phase, the  $R^2$  values for each model ranged from 0.39 to 0.69 and the RMSE values varied from 0.2 to 0.3. During the test phase, the ranges for RMSE and  $R^2$  were 0.038–0.044 and 0.33–0.55, respectively. As with the stomatal conductance to water vapor and net assimilation rate, REGST and CNN-REGST showed relatively better results in predicting the quantum yield of PSII in comparison to other models.

For transpiration rate, the results from training models had RMSE and  $R^2$  values between 0.91 and 1.17 and 0.60 and 0.77, respectively. REGST model outperformed the rest of the training algorithms (Table 3(iv)). The model performances in the test dataset varied with RMSE of 1.19–1.41 and  $R^2$  of 0.38–0.67. Again, REGST and CNN-REGST had very similar RMSE and  $R^2$  metrics in their test performances for the prediction of transpiration rate.

### 3.4. Important Hyperspectral Features in Modeling Algorithm

To identify the important spectral features and their rank contributing to REGST model performance, the permutation importance was calculated. The feature importance score of each wavelength in the test dataset was retrieved for the REGST model. Figure 6 illustrates feature important values estimated for each of the response variables, i.e., net assimilation rate, stomatal conductance to water vapor, quantum yield of PSII, and transpiration rate.



**Figure 6.** Feature importance score assigned for each wavelength by the REGST model. Net carbon assimilation rate,  $\mu\text{mol m}^{-2} \text{s}^{-1}$  (A, green); stomatal conductance to water vapor,  $\text{mol m}^{-2} \text{s}^{-1}$  (gsw, purple); quantum yield of PSII, ( $\phi$  PSII, mauve), and transpiration,  $\text{mmol m}^{-2} \text{s}^{-1}$  (E, blue).

Considering all the physiological traits, the first important spectral region was consistently occurring in the 490–510 nm wavelength. For instance, the importance value was around 0.4 for stomatal conductance to water vapor and 0.25 for quantum yield of PSII. The other major region is around the 700–750 nm region, which consistently occurred as the most important predictor variable. This applies for stomatal conductance to water vapor (~0.4), quantum yield of PSII (~0.25), and transpiration rate (~0.30). Finally, the third important region is around 900–950 nm. In Figure 6, it can be observed that the wavelengths under this range had the highest importance scores in all cases.

## 4. Discussion

### 4.1. Rootstock Has Significant Impact on Physiology of Scion

Significant variations in net assimilation rate, stomatal conductance to water vapor, and transpiration rate were identified during the evaluation of the ‘Marquette’ scion physiology grafted to the different rootstocks. The analyses showed that ‘Marquette’ on 3309C, 1103P, and Freedom rootstocks exhibited notably higher levels of net assimilation rate, transpiration rate, and stomatal conductance to water vapor. The performance rate of the ‘Marquette’ scion grafted on 5C and SO4 was much lower for all physiological characteristics, and visual observations indicated that the scion was less vigorous in these graft combinations. Like this finding, Bica et al. (2000) [60] observed a considerable impact of rootstocks on various physiological parameters. The photosynthetic activity, quantum yield, stomatal conductance, and chlorophyll content of Chardonnay vines grafted onto SO4 rootstock were found to be reduced compared to those grafted onto 1103P rootstock [60]. In this study, it is noteworthy to mention that the observed variations in the results are solely attributable to rootstock genotype factors, as all external environmental factors were the same. Both SO4 and 5C rootstocks were derived from parental crosses between *V. berlandieri* and *V. riparia*. In contrast, the other rootstocks have different parentages, 3309C (*V. riparia* × *V. rupestris*), 1103P (*V. berlandieri* × *V. rupestris*), and Freedom (*V. champinii* × (*V. colonis* × *V. othello*)). One potential reason that may contribute to the observed variations in vigor and performance of the measured physiological parameters between SO4 and 5C and 3309C, 1103P and Freedom may be the presence of *V. rupestris* in the later rootstock’s genetic composition. Pou et al. (2022) [13] conducted a study that investigated the effects of four different rootstock types on the *V. vinifera* ‘Tempranillo’ cultivar as a scion for different physiological parameters. The study found that the photosynthesis and stomatal conductance were comparatively lower in the rootstock with a *V. riparia* background when compared to *V. rupestris*. In this case, although there is *V. riparia* in all the pedigrees, the contrast may be related to either the *V. rupestris* or the *V. berlandieri*. However, many studies have reported that rootstocks containing *V. rupestris* in their genetic makeup have enhanced abilities in root water uptake, transport capacity, and transpiration compared to other rootstocks [61–63]. The photosynthetic performance as influenced by rootstocks shown here aligns with previous suggestions that rootstock selection can influence the physiological characteristics of the scion.

### 4.2. Aerial Hyperspectral Data Have a Positive Correlation with Actual Ground-Truth Physiology Parameters of Vines

The measured physiological phenotypes showed that the different grafted vines had significant variation in their performances. Physiological traits like carbon assimilation, stomatal conductance, and transpiration are often associated with leaf attributes like water content [64], mesophyll cell structure/distribution [65], and cell wall composition [66]. Variation in these leaf anatomical and biochemical attributes is specific to variations in spectral reflectance at different wavelengths [67,68].

Two highly correlated regions between physiological traits and wavelengths (Figure 5) were observed (520–565 nm, and 730–790 nm). In Section 3.4, the region around 900–950 nm was identified as the most important feature for modeling analysis. All these wavelengths have been linked with structural and physiological traits in perennials. The 520–565 nm



wavelengths, corresponding to green spectral region, is closely linked to photosynthetic-light-use dynamics with its sensitivity to changes in xanthophyll pigment cycle [69–72]. The PRI (photochemical reflective index) derived from this region has been shown to be sensitive to changes in photosynthesis and found to have strong positive correlation with field measured assimilation rates in many studies [73,74].

The 730–790 nm range corresponds to the red-edge region (680–750 nm) and near-infrared region (700–1400 nm). These regions are traditionally used to calculate different vegetative indices like the Green Normalized Difference Vegetation Index (GNDVI) and Anthocyanin [75]. Several studies show significant positive correlation between this spectral region and physiological traits like stomatal conductance and leaf water potential in grapevines [29,30]. The spectral range of 900–950 nm, which is the near-infrared region was identified as an important feature for the prediction of physiological parameters in Section 3.4. Many studies indicate the spectra around 950 nm is associated with plant internal leaf mesophyll characteristics like cell wall and stomatal conductance [76–78].

#### 4.3. Ensemble Stacked Regression (REGST) Model Outperformed Individual Base Models

All the individual base models in this study were designed as feature extractors to target multicollinearity and perform dimension reduction of hyperspectral wavelengths. Depending on different scenarios, statistical algorithms based on latent variables (PLSR and PCR) as well as regularization techniques (Lasso, RR, ENET) had different prediction accuracies. One common issue with these models was their inefficacy to precisely predict smaller values of each physiological trait (Supplementary File Figures S1–S4). Comparing base model metrics, RR had the lowest test RMSE; however, PLSR outperformed the rest of the base models for stomatal conductance. Given that each algorithm possesses varying predictive capabilities, it is essential to explore the possibility of synergistically leveraging the strengths of each predictive model. Based on this concept, few studies have utilized ensemble stacking in remote sensing applications [44,79–81]. Our study employed the technique of aggregating diverse regression models with the hypothesis that it would increase predictive performance through ensemble stacking regression relative to the base individual algorithms. The REGST model was found to increase  $R^2$  by 0.03–0.1 compared to base individual models. One important finding was that ensemble stacked regression significantly increased prediction accuracy in our scenario where the number of ground-truth samples was much smaller than the hyperspectral features utilized as predictor variables. The REGST model created a new and larger set of training data for the meta-learner by combining predictions from base models, thereby overcoming data size limitations. The implementation of the stacking ensemble model is shown to boost prediction accuracy in other studies. Fu et al., 2019 [44], in their paper estimating tobacco (*Nicotiana tabacum*) photosynthetic capacities using stacking regression ensemble approach, found an increase in  $R^2$  of 0.1 (0.08) compared to all individual base models. They used artificial neural network (ANN), support vector machine (SVM), least absolute shrinkage and selection operator (Lasso), random forest (RF), gaussian process regression (GP), and partial least square regression (PLSR) models as base models to predict  $V_{\text{cmax}}$  and  $J_{\text{max}}$  from hyperspectral signal retrieved from tobacco genotypes. Similarly, in maize using a proximal hyperspectral sensor to predict leaf chlorophyll content, Huang et al. [81] observed a significant increase in the  $R^2$  value using stacking regressor as compared to individual base models (support vector regression (SVR), back propagation neural network (BPNN), and PLSR). Similarly, our study showed that stacking regression is a powerful tool for predicting physiological traits in grapevines. Nevertheless, the predictive performance of REGST could still be refined through the incorporation of more ML methods, deep learning algorithms as base models, or feature engineering of input variables.

#### 4.4. Both REGST and Hybrid CNN-REGST Models Had Similar Predictive Performances

Very few findings have been published in the remote sensing genre regarding the CNN-REGST model. One study in wheat [82], using a hybrid 1D CNN and ensemble

model, shows an overall increase in  $R^2$  from 0.74 to 0.79 while predicting  $V_{\text{cmax}}$ ,  $J_{\text{max}}$ , and the assimilation rate. Despite extensive searches of the literature, there seems to be a very limited number of studies that explore or demonstrate the application of such hybrid models for other physiological trait predictions. Therefore, this study applied feature engineering using a CNN algorithm to select input features for REGST models to improve model accuracy. After the selection of optimal layers, kernel sizes, and hyperparameters, a total of 80 learned features were extracted from flattened layers as input for REGST. While comparing the performance of REGST with or without feature engineering using CNN, the model metric  $R^2$  retrieved for each physiological trait was either identical or varied by 0.01. For the net assimilation rate, the REGST model had the lowest test RMSE (3.51) as compared to CNN-REGST (3.55). In the case of stomatal conductance to water vapor, CNN-REGST had a slightly lower test RMSE error (0.078) compared to REGST (0.083). Both REGST and CNN-REGST were considered the best models for their performance in estimating the transpiration rate and quantum yield of PSII. The common performance pattern in the prediction of each physiological trait suggested that both the models share similarity in features used for training the respective algorithms. In contrast, a study on ground-based solar radiation prediction indicates that CNN-REGST has the least relative root mean square error in six study locations followed by the CNN and REGST models [83]. In their study, the CNN model alone has better performance and, when combined with REGST (CNN-REGST), hugely improves the prediction accuracy in the hybrid model. When we evaluated the performance of CNN alone in this study, it had higher RMSE values compared to all other model metrics. One possible reason is due to the size of the dataset. Smaller datasets present insufficient diverse examples for CNN to learn meaningful patterns. This causes poor generalization in the training dataset, eventually leading to poor performance in the test dataset [84,85]. As a result, combining CNN with REGST did not improve the prediction performance as much as expected in this study. To further explore the potential of CNN-REGST compared to the REGST model, further experimentation with an alternative large dataset is needed. Likewise, while the proposed CNN-REGST model showed promising results, their use in real-world vineyard management requires further validation and refining to provide robustness in a variety of climatic circumstances. Furthermore, there are still challenges that must be resolved, such as the computational demands of processing extensive hyperspectral datasets and the necessity of substantial training data.

Both the modeling approaches used in this study, REGST and CNN-REGST, excel at extracting significant features. REGST reduces feature dimensionality through base algorithms, while CNN-REGST automates feature engineering by capturing local dependencies. However, the two models provide similar results in this study, indicating either model can be used effectively. The analyses in Sections 3.2 and 3.4 provide valuable insights into how specific wavelengths can be effectively utilized in future physiological research. When these specific wavelengths are integrated into machine learning algorithms, they not only simplify the model complexity but also can enhance its generalization capabilities across diverse datasets. This streamlined approach ensures that each selected wavelength contributes meaningfully to the model's predictive power, fostering robustness and efficiency in subsequent studies.

## 5. Conclusions

The complex structural and biophysical attributes in grapevines and other woody perennials make direct assessment of their physiology very challenging. This study integrated hyperspectral aerial remote sensing with an AI-based modeling approach to indirectly measure different physiological traits. Net assimilation rate, stomatal conductance to water vapor, quantum yield of PSII, and transpiration rate were measured as ground-truth physiological indicators. Spectral data from an aerial hyperspectral sensor were analyzed using various ML algorithms, ensemble stacked regression (REGST) and hybrid of CNN, and ensemble stacked regression (CNN-REGST) algorithms. Results

from the different approaches indicated that REGST had better prediction performance compared to individual base models with an increase in  $R^2$  by nearly 0.1. Nevertheless, the comparative assessment of the hybrid CNN-REGST and REGST showed very similar predictive performance, although this could be due to the issue of limited data availability. The overall findings indicated that hyperspectral aerial remote sensing has high potential as a robust high-throughput technique in quantifying plant functional traits. In addition, the modeling algorithm assessments over a two-year data period indicate that the integration of AI-based powerful techniques improves potential indirect measures to monitor plant physiological function. The REGST and CNN-REGST models' capacity to process and evaluate hyperspectral images enables a non-invasive, quick, and large-scale monitoring tool that outperforms standard visual inspection approaches. In addition, integrating specific wavelengths, which are highly correlated with the physiological traits, into machine learning algorithms simplifies the model complexity and can potentially provide insights into characteristics such as photosynthetic light use, stomatal conductance, leaf water potential, and leaf mesophyll. The results of the study offer viticulturists actionable insights, enabling the optimization of monitoring activities and the early detection of stress conditions, which could potentially reduce costs in vineyard management. This effort furthers the development of precision agriculture practices in viticulture, with the goal of improving vineyard profitability and sustainability by bridging the gap between theoretical research and actual application.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/rs16142626/s1>, Figure S1: Comparison of predictive accuracy across different regression models for net assimilation rate. Figure S2: Comparison of predictive accuracy across different regression models for stomatal conductance to water vapor. Figure S3: Comparison of predictive accuracy across different regression models for quantum yield of PSII. Figure S4: Comparison of predictive accuracy across different regression models for Transpiration rate (E). Table S1. Field Map. Table S2. Hyperparameter tuning grid for various machine learning and deep learning models.

**Author Contributions:** Conceptualization, P.S., R.V.-D. and A.F.; methodology, P.S. and R.V.-D.; Software, R.V.-D.; formal analysis, P.S.; resources, A.F.; data curation, P.S.; writing—original draft preparation, P.S.; writing—review and editing, A.F.; supervision, A.F.; project administration, A.F.; funding acquisition, A.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Science Foundation (NSF) grant number 154685, and the South Dakota Agricultural Experiment Station, Hatch Project No. SD00H765.

**Data Availability Statement:** Raw data are available through Sharma, P. and Fennell, A. Reflectance data for predicting grapevine physiological parameters. OpenPrairie.sdstate.edu, accessed on 8 July 2024.

**Acknowledgments:** We acknowledge the South Dakota State University Research Computing members for assistance in High-Performance Computing needs, debugging, and managing large memory/GPU requirements.

**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.

## References

1. Martínez de Toda, F.; Balda, P. Delaying berry ripening through manipulating leaf area to fruit ratio. *J. Grapevine Res.* **2013**, *52*, 171–176.
2. Martinez de Toda, F.; Sancha, J.C.; Balda, P. Reducing the sugar and pH of the grape (*Vitis vinifera* L. cvs. 'Grenache' and 'Tempranillo') through a single shoot trimming. *S. Afr. J. Enol. Vitic.* **2013**, *34*, 246–251.
3. Keller, M.; Tarara, J.; Mills, L. Spring temperatures alter reproductive development in grapevines. *Aust. J. Grape Wine Res.* **2010**, *16*, 445–454. [[CrossRef](#)]
4. Parker, A.K.; De Cortazar-Atauri, I.G.; Van Leeuwen, C.; Chuine, I. General phenological model to characterise the timing of flowering and veraison of *Vitis vinifera* L. *Aust. J. Grape Wine Res.* **2011**, *17*, 206–216. [[CrossRef](#)]

5. Ruml, M.; Korać, N.; Vujadinović, M.; Vuković, A.; Ivanišević, D. Response of grapevine phenology to recent temperature change and variability in the wine-producing area of Sremski Karlovci, Serbia. *J. Agric. Sci.* **2016**, *154*, 186–206. [\[CrossRef\]](#)
6. Schultz, H.; Hofmann, M. The ups and downs of environmental impact on grapevines: Future challenges in temperate viticulture. In *Grapevine in a Changing Environment: A Molecular and Ecophysiological Perspective*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2015; pp. 18–37.
7. Arrizabalaga-Arriazu, M.; Morales, F.; Irigoyen, J.J.; Hilbert, G.; Pascual, I. Growth performance and carbon partitioning of grapevine Tempranillo clones under simulated climate change scenarios: Elevated CO<sub>2</sub> and temperature. *J. Plant Physiol.* **2020**, *252*, 153226. [\[CrossRef\]](#)
8. Galet, P. Phylloxera galls on *Vitis vinifera* L. [Grapes]. *Prog. Agric. Vitic.* **1983**, *100*, 155–162.
9. Ordish, G. *The Great Wine Blight*; J.M. Dent & Sons Ltd.: London, UK, 1972.
10. Pouget, R. *Histoire de la Lutte Contre le Phylloxéra de la Vigne en France*; INRA: Paris, France, 1990; pp. 1–168.
11. Corso, M.; Bonghi, C. Grapevine rootstock effects on abiotic stress tolerance. *Plant Sci. Today* **2014**, *1*, 108–113. [\[CrossRef\]](#)
12. Warschefsky, E.J.; Klein, L.L.; Frank, M.H.; Chitwood, D.H.; Londo, J.P.; von Wettberg, E.J.; Miller, A.J. Rootstocks: Diversity, domestication, and impacts on shoot phenotypes. *Trends Plant Sci.* **2016**, *21*, 418–437. [\[CrossRef\]](#)
13. Pou, A.; Rivacoba, L.; Portu, J.; Mairata, A.; Labarga, D.; García-Escudero, E.; Martín, I. How Rootstocks Impact the Scion Vigour and Vine Performance of *Vitis vinifera* L. cv. Tempranillo. *Aust. J. Grape Wine Res.* **2022**, *2022*, 9871347. [\[CrossRef\]](#)
14. Mantilla, S.M.O.; Collins, C.; Iland, P.G.; Kidman, C.M.; Ristic, R.; Boss, P.K.; Jordans, C.; Bastian, S.E. Shiraz (*Vitis vinifera* L.) berry and wine sensory profiles and composition are modulated by rootstocks. *Am. J. Enol. Vitic.* **2018**, *69*, 32–44. [\[CrossRef\]](#)
15. Keller, M. Developmental physiology. In *The Science of Grapevines: Anatomy and Physiology*; Academic Press: Cambridge, MA, USA, 2010; pp. 169–225.
16. Padgett-Johnson, M.; Williams, L.; Walker, M.A. The influence of *Vitis riparia* rootstock on water relations and gas exchange of *Vitis vinifera* cv. Carignane scion under non-irrigated conditions. *Am. J. Enol. Vitic.* **2000**, *51*, 137–143. [\[CrossRef\]](#)
17. Soar, C.J.; Dry, P.R.; Loveys, B. Scion photosynthesis and leaf gas exchange in *Vitis vinifera* L. cv. Shiraz: Mediation of rootstock effects via xylem sap ABA. *Aust. J. Grape Wine Res.* **2006**, *12*, 82–96. [\[CrossRef\]](#)
18. Marguerit, E.; Brendel, O.; Lebon, E.; Van Leeuwen, C.; Ollat, N. Rootstock control of scion transpiration and its acclimation to water deficit are controlled by different genes. *New Phytol.* **2012**, *194*, 416–429. [\[CrossRef\]](#)
19. Suarez, D.L.; Celis, N.; Anderson, R.G.; Sandhu, D. Grape Rootstock Response to Salinity, Water and Combined Salinity and Water Stresses. *Agronomy* **2019**, *9*, 321. [\[CrossRef\]](#)
20. Domingues Neto, F.J.; Pimentel Junior, A.; Modesto, L.R.; Moura, M.F.; Putti, F.F.; Boaro, C.S.F.; Ono, E.O.; Rodrigues, J.D.; Tecchio, M.A. Photosynthesis, Biochemical and Yield Performance of Grapevine Hybrids in Two Rootstock and Trellis Height. *Horticulturae* **2023**, *9*, 596. [\[CrossRef\]](#)
21. Prinsi, B.; Simeoni, F.; Galbiati, M.; Meggio, F.; Tonelli, C.; Scienza, A.; Espen, L. Grapevine rootstocks differently affect physiological and molecular responses of the scion under water deficit condition. *Agronomy* **2021**, *11*, 289. [\[CrossRef\]](#)
22. Ghule, V.; Zagade, P.; Bhor, V.; Somkuwar, R. Rootstock affects graft success growth and physiological parameters of grape varieties (*Vitis vinifera* L.). *Int. J. Curr. Microbiol. App. Sci* **2019**, *8*, 799–805. [\[CrossRef\]](#)
23. Edwards, E.; Betts, A.; Clingeffer, P.; Walker, R. Rootstock-conferred traits affect the water use efficiency of fruit production in Shiraz. *Aust. J. Grape Wine Res.* **2022**, *28*, 316–327. [\[CrossRef\]](#)
24. Bianchi, D.; Ricciardi, V.; Pozzoli, C.; Grossi, D.; Caramanico, L.; Pindo, M.; Stefani, E.; Cestaro, A.; Brancadoro, L.; De Lorenzis, G. Physiological and Transcriptomic Evaluation of Drought Effect on Own-Rooted and Grafted Grapevine Rootstock (1103P and 101-14MGt). *Plants* **2023**, *12*, 1080. [\[CrossRef\]](#)
25. Peccoux, A.; Loveys, B.; Zhu, J.; Gambetta, G.A.; Delrot, S.; Vivin, P.; Schultz, H.R.; Ollat, N.; Dai, Z. Dissecting the rootstock control of scion transpiration using model-assisted analyses in grapevine. *Tree Physiol.* **2017**, *38*, 1026–1040. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Maimaitiyiming, M.; Sagan, V.; Sidike, P.; Maimaitijiang, M.; Miller, A.J.; Kwasniewski, M. Leveraging very-high spatial resolution hyperspectral and thermal UAV imageries for characterizing diurnal indicators of grapevine physiology. *Remote Sens.* **2020**, *12*, 3216. [\[CrossRef\]](#)
27. Yang, Z.; Tian, J.; Wang, Z.; Feng, K. Monitoring the photosynthetic performance of grape leaves using a hyperspectral-based machine learning model. *Eur. J. Agron.* **2022**, *140*, 126589. [\[CrossRef\]](#)
28. Zarco-Tejada, P.J.; Catalina, A.; González, M.; Martín, P. Relationships between net photosynthesis and steady-state chlorophyll fluorescence retrieved from airborne hyperspectral imagery. *Remote Sens. Environ.* **2013**, *136*, 247–258. [\[CrossRef\]](#)
29. Pôças, I.; Rodrigues, A.; Gonçalves, S.; Costa, P.M.; Gonçalves, I.; Pereira, L.S.; Cunha, M. Predicting grapevine water status based on hyperspectral reflectance vegetation indices. *Remote Sens.* **2015**, *7*, 16460–16479. [\[CrossRef\]](#)
30. Rodríguez-Pérez, J.R.; Riaño, D.; Carlisle, E.; Ustin, S.; Smart, D.R. Evaluation of hyperspectral reflectance indexes to detect grapevine water status in vineyards. *Am. J. Enol. Vitic.* **2007**, *58*, 302–317. [\[CrossRef\]](#)
31. Mariotto, I.; Thenkabail, P.S.; Huete, A.; Slonecker, E.T.; Platonov, A. Hyperspectral versus multispectral crop-productivity modeling and type discrimination for the HypSIIRI mission. *Remote Sens. Environ.* **2013**, *139*, 291–305. [\[CrossRef\]](#)
32. Anece, I.; Thenkabail, P. Accuracies achieved in classifying five leading world crop types and their growth stages using optimal earth observing-1 hyperion hyperspectral narrowbands on google earth engine. *Remote Sens.* **2018**, *10*, 2027. [\[CrossRef\]](#)
33. Fu, P.; Meacham-Hensold, K.; Guan, K.; Wu, J.; Bernacchi, C. Estimating photosynthetic traits from reflectance spectra: A synthesis of spectral indices, numerical inversion, and partial least square regression. *Plant Cell Environ.* **2020**, *43*, 1241–1258. [\[CrossRef\]](#)



34. Serbin, S.P.; Dillaway, D.N.; Kruger, E.L.; Townsend, P.A. Leaf optical properties reflect variation in photosynthetic metabolism and its sensitivity to temperature. *J. Exp. Bot.* **2012**, *63*, 489–502. [CrossRef]
35. Wold, H. Systems under indirect observation using, P.L.S. In *A Second Generation of Multivariate Analysis: Methods*; Praeger: Westport, CT, USA, 1982.
36. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [CrossRef]
37. Geladi, P.; Kowalski, B.R. Partial least-squares regression: A tutorial. *Anal. Chim. Acta* **1986**, *185*, 1–17. [CrossRef]
38. Jolliffe, I.T. A note on the use of principal components in regression. *J. R. Stat. Soc. Ser. C Appl. Stat.* **1982**, *31*, 300–303. [CrossRef]
39. Gomes, V.; Rendall, R.; Reis, M.S.; Mendes-Ferreira, A.; Melo-Pinto, P. Determination of sugar, pH, and anthocyanin contents in port wine grape berries through hyperspectral imaging: An extensive comparison of linear and non-linear predictive methods. *Appl. Sci.* **2021**, *11*, 10319. [CrossRef]
40. Yang, Y.; Nan, R.; Mi, T.; Song, Y.; Shi, F.; Liu, X.; Wang, Y.; Sun, F.; Xi, Y.; Zhang, C. Rapid and nondestructive evaluation of wheat chlorophyll under drought stress using hyperspectral imaging. *Int. J. Mol. Sci.* **2023**, *24*, 5825. [CrossRef] [PubMed]
41. Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1996**, *58*, 267–288. [CrossRef]
42. Marquardt, D.W.; Snee, R.D. Ridge regression in practice. *Am. Stat.* **1975**, *29*, 3–20. [CrossRef]
43. Zou, H.; Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2005**, *67*, 301–320. [CrossRef]
44. Fu, P.; Meacham-Hensold, K.; Guan, K.; Bernacchi, C.J. Hyperspectral leaf reflectance as proxy for photosynthetic capacities: An ensemble approach based on multiple machine learning algorithms. *Front. Plant Sci.* **2019**, *10*, 454448. [CrossRef]
45. Wolpert, D.H. Stacked generalization. *Neural Netw.* **1992**, *5*, 241–259. [CrossRef]
46. Unlarsen, M.F.; Sonmez, M.E.; Aslan, M.F.; Demir, B.; Aydin, N.; Sabanci, K.; Ropelewska, E. CNN-SVM hybrid model for varietal classification of wheat based on bulk samples. *Eur. Food Res. Technol.* **2022**, *248*, 2043–2052. [CrossRef]
47. Tanwar, V.; Lamba, S.; Sharma, B. Deep learning-based hybrid model for severity prediction of leaf smut sugarcane infection. In Proceedings of the 2023 Third International Conference on Artificial Intelligence and Smart Energy (ICAIS), Coimbatore, India, 2–4 February 2023; IEEE: Piscataway, NJ, USA, 2023.
48. Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; IEEE: Piscataway, NJ, USA, 2017.
49. Gavrishchaka, V.; Yang, Z.; Miao, R.; Senyukova, O. Advantages of hybrid deep learning frameworks in applications with limited data. *Int. J. Mach. Learn. Comput.* **2018**, *8*, 549–558.
50. Smith, G.M.; Milton, E.J. The use of the empirical line method to calibrate remotely sensed data to reflectance. *Int. J. Remote Sens.* **1999**, *20*, 2653–2662. [CrossRef]
51. Gillies, S. *Rasterio Documentation*; MapBox: San Francisco, CA, USA, 2019; Volume 23.
52. Press, W.H.; Teukolsky, S.A. Savitzky-Golay smoothing filters. *Comput. Phys.* **1990**, *4*, 669–672. [CrossRef]
53. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [CrossRef] [PubMed]
54. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
55. Fushiki, T. Estimation of prediction error by using K-fold cross-validation. *Stat. Comput.* **2011**, *21*, 137–146. [CrossRef]
56. Bergstra, J.; Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
57. Berrar, D. Cross-Validation. In *Reference Module in Life Sciences Encyclopedia of Bioinformatics and Computational Biology*; Elsevier: Amsterdam, The Netherlands, 2019.
58. Altmann, A.; Tolosi, L.; Sander, O.; Lengauer, T. Permutation importance: A corrected feature importance measure. *Bioinformatics* **2010**, *26*, 1340–1347. [CrossRef]
59. Raschka, S. MLxtend: A Python Library for Machine Learning Extensions. Available online: <https://sebastianraschka.com/pdf/software/mlxtend-latest.pdf> (accessed on 1 May 2023).
60. Bica, D.; Gay, G.; Morando, A.; Soave, E. Effects of rootstock and *Vitis vinifera* genotype on photosynthetic parameters. In Proceedings of the V International Symposium on Grapevine Physiology 526, Jerusalem, Israel, 25–30 May 1997.
61. Alsina, M.M.; Smart, D.R.; Bauerle, T.; De Herralde, F.; Biel, C.; Stockert, C.; Negron, C.; Save, R. Seasonal changes of whole root system conductance by a drought-tolerant grape root system. *J. Exp. Bot.* **2011**, *62*, 99–109. [CrossRef]
62. Romero, P.; Botía, P.; Navarro, J.M. Selecting rootstocks to improve vine performance and vineyard sustainability in deficit irrigated Monastrell grapevines under semiarid conditions. *Agric. Water Manag.* **2018**, *209*, 73–93. [CrossRef]
63. Lovisolo, C.; Perrone, I.; Carra, A.; Ferrandino, A.; Flexas, J.; Medrano, H.; Schubert, A. Drought-induced changes in development and function of grapevine (*Vitis* spp.) organs and in their hydraulic and non-hydraulic interactions at the whole-plant level: A physiological and molecular update. *Funct. Plant Biol.* **2010**, *37*, 98–116. [CrossRef]
64. Curran, P.J. Remote sensing of foliar chemistry. *Remote Sens. Environ.* **1989**, *30*, 271–278. [CrossRef]



65. Lehmeier, C.; Pajor, R.; Lundgren, M.R.; Mathers, A.; Sloan, J.; Bauch, M.; Mitchell, A.; Bellasio, C.; Green, A.; Bouyer, D. Cell density and airspace patterning in the leaf can be manipulated to increase leaf photosynthetic capacity. *Plant J.* **2017**, *92*, 981–994. [[CrossRef](#)] [[PubMed](#)]
66. Peñuelas, J.; Filella, I. Visible and near-infrared reflectance techniques for diagnosing plant physiological status. *Trends Plant Sci.* **1998**, *3*, 151–156. [[CrossRef](#)]
67. Liu, L.; Wang, J.; Huang, W.; Zhao, C.; Zhang, B.; Tong, Q. Estimating winter wheat plant water content using red edge parameters. *Int. J. Remote Sens.* **2004**, *25*, 3331–3342. [[CrossRef](#)]
68. Tilling, A.K.; O’Leary, G.J.; Ferwerda, J.G.; Jones, S.D.; Fitzgerald, G.J.; Rodriguez, D.; Belford, R. Remote sensing of nitrogen and water stress in wheat. *Field Crops Res.* **2007**, *104*, 77–85. [[CrossRef](#)]
69. Gamon, J.; Penuelas, J.; Field, C. A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sens. Environ.* **1992**, *41*, 35–44. [[CrossRef](#)]
70. Gamon, J.; Serrano, L.; Surfus, J. The photochemical reflectance index: An optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels. *Oecologia* **1997**, *112*, 492–501. [[CrossRef](#)]
71. Peñuelas, J.; Filella, I.; Gamon, J.A. Assessment of photosynthetic radiation-use efficiency with spectral reflectance. *New Phytol.* **1995**, *131*, 291–296. [[CrossRef](#)]
72. Peñuelas, J.; Llusia, J.; Pinol, J.; Filella, I. Photochemical reflectance index and leaf photosynthetic radiation-use-efficiency assessment in Mediterranean trees. *Int. J. Remote Sens.* **1997**, *18*, 2863–2868. [[CrossRef](#)]
73. Suarez, L.; González-Dugo, V.; Camino, C.; Hornero, A.; Zarco-Tejada, P.J. Physical model inversion of the green spectral region to track assimilation rate in almond trees with an airborne nano-hyperspectral imager. *Remote Sens. Environ.* **2021**, *252*, 112147. [[CrossRef](#)]
74. Hikosaka, K.; Tsujimoto, K. Linking remote sensing parameters to CO<sub>2</sub> assimilation rates at a leaf scale. *J. Plant Res.* **2021**, *134*, 695–711. [[CrossRef](#)] [[PubMed](#)]
75. Gitelson, A.A.; Merzlyak, M.N.; Chivkunova, O.B. Optical properties and nondestructive estimation of anthocyanin content in plant leaves. *Photochem. Photobiol.* **2001**, *74*, 38–45. [[CrossRef](#)] [[PubMed](#)]
76. Berger, B.; Parent, B.; Tester, M. High-throughput shoot imaging to study drought responses. *J. Exp. Bot.* **2010**, *61*, 3519–3528. [[CrossRef](#)] [[PubMed](#)]
77. Guo, J.-T.; Yang, D.-C.; Guan, Z.; He, Y.-H. Chlorophyll-catalyzed visible-light-mediated synthesis of tetrahydroquinolines from N,N-dimethylanilines and maleimides. *J. Org. Chem.* **2017**, *82*, 1888–1894. [[CrossRef](#)] [[PubMed](#)]
78. Slaton, M.R.; Hunt, E.R., Jr.; Smith, W.K. Estimating near-infrared leaf reflectance from leaf structural characteristics. *Am. J. Bot.* **2001**, *88*, 278–284. [[CrossRef](#)] [[PubMed](#)]
79. Clinton, N.; Yu, L.; Gong, P. Geographic stacking: Decision fusion to increase global land cover map accuracy. *ISPRS J. Photogramm. Remote Sens.* **2015**, *103*, 57–65. [[CrossRef](#)]
80. Healey, S.P.; Cohen, W.B.; Yang, Z.; Brewer, C.K.; Brooks, E.B.; Gorelick, N.; Hernandez, A.J.; Huang, C.; Hughes, M.J.; Kennedy, R.E. Mapping forest change using stacked generalization: An ensemble approach. *Remote Sens. Environ.* **2018**, *204*, 717–728. [[CrossRef](#)]
81. Huang, X.; Guan, H.; Bo, L.; Xu, Z.; Mao, X. Hyperspectral proximal sensing of leaf chlorophyll content of spring maize based on a hybrid of physically based modelling and ensemble stacking. *Comput. Electron. Agric.* **2023**, *208*, 107745. [[CrossRef](#)]
82. Furbank, R.T.; Silva-Perez, V.; Evans, J.R.; Condon, A.G.; Estavillo, G.M.; He, W.; Newman, S.; Poiré, R.; Hall, A.; He, Z. Wheat physiology predictor: Predicting physiological traits in wheat from hyperspectral reflectance measurements using deep learning. *Plant Methods* **2021**, *17*, 108. [[CrossRef](#)]
83. Ghimire, S.; Nguyen-Huy, T.; Deo, R.C.; Casillas-Perez, D.; Salcedo-Sanz, S. Efficient daily solar radiation prediction with deep learning 4-phase convolutional neural network, dual stage stacked regression and support vector machine CNN-REGST hybrid model. *Sustain. Mater. Technol.* **2022**, *32*, e00429. [[CrossRef](#)]
84. Kamilaris, A.; Prenafeta-Boldú, F.X. A review of the use of convolutional neural networks in agriculture. *J. Agric. Sci.* **2018**, *156*, 312–322. [[CrossRef](#)]
85. Alzubaidi, L.; Bai, J.; Al-Sabaawi, A.; Santamaría, J.; Albahri, A.; Al-dabbagh, B.S.N.; Fadhel, M.A.; Manoufali, M.; Zhang, J.; Al-Timemy, A.H. A survey on deep learning tools dealing with data scarcity: Definitions, challenges, solutions, tips, and applications. *J. Big Data* **2023**, *10*, 46. [[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.