



Article

Predicting Canopy Chlorophyll Content in Sugarcane Crops Using Machine Learning Algorithms and Spectral Vegetation Indices Derived from UAV Multispectral Imagery

Amarasingam Narmilan ^{1,2,*} , Felipe Gonzalez ¹ , Arachchige Surantha Ashan Salgadoe ³ , Unupen Widanelage Lahiru Madhushanka Kumarasiri ⁴ , Hettiarachchige Asiri Sampageeth Weerasinghe ⁴ and Buddhika Rasanjana Kulasekara ⁴

- ¹ School of Electrical Engineering and Robotics, Faculty of Engineering, Queensland University of Technology (QUT), 2 George Street, Brisbane, QLD 4000, Australia; felipe.gonzalez@qut.edu.au
² Department of Biosystems Technology, Faculty of Technology, South Eastern University of Sri Lanka, University Park, Oluvil 32360, Sri Lanka
³ Department of Horticulture and Landscape Gardening, Wayamba University of Sri Lanka, Makandura, Gonawila 60170, Sri Lanka; surantha@wyb.ac.lk
⁴ Division of Crop Nutrition, Sugarcane Research Institute, Dakunu Ela Road, Udawalawe 70190, Sri Lanka; kumarasiri99@sugarres.lk (U.W.L.M.K.); asiriwee@sugarres.lk (H.A.S.W.); kulasekara@sugarres.lk (B.R.K.)
* Correspondence: narmilan.amarasingam@hdr.qut.edu.au or narmilan@seu.ac.lk



Citation: Narmilan, A.; Gonzalez, F.; Salgadoe, A.S.A.; Kumarasiri, U.W.L.M.; Weerasinghe, H.A.S.; Kulasekara, B.R. Predicting Canopy Chlorophyll Content in Sugarcane Crops Using Machine Learning Algorithms and Spectral Vegetation Indices Derived from UAV Multispectral Imagery. *Remote Sens.* **2022**, *14*, 1140. <https://doi.org/10.3390/rs14051140>

Academic Editors: Thomas Alexandridis, Tomáš Řezník, Abdul M. Mouazen and Guo Huadong

Received: 22 January 2022
Accepted: 23 February 2022
Published: 25 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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/>).

Abstract: The use of satellite-based Remote Sensing (RS) is a well-developed field of research. RS techniques have been successfully utilized to evaluate the chlorophyll content for the monitoring of sugarcane crops. This research provides a new framework for inferring the chlorophyll content in sugarcane crops at the canopy level using unmanned aerial vehicles (UAVs) and spectral vegetation indices processed with multiple machine learning algorithms. Studies were conducted in a sugarcane field located in Sugarcane Research Institute (SRI, Uda Walawe, Sri Lanka), with various fertilizer applications over the entire growing season from 2020 to 2021. An UAV with multispectral camera was used to collect the aerial images to generate the vegetation indices. Ground measurements of leaf chlorophyll were used as indications for fertilizer status in the sugarcane field. Different machine learning (ML) algorithms were used ground-truthing data of chlorophyll content and spectral vegetation indices to forecast sugarcane chlorophyll content. Several machine learning algorithms such as MLR, RF, DT, SVR, XGB, KNN and ANN were applied in two ways: before feature selection (BFS) by training the algorithms with all twenty-four (24) vegetation indices with five (05) spectral bands and after feature selection (AFS) by training algorithms with fifteen (15) vegetation indices. All the algorithms with both BFS and AFS methods were compared with an estimated coefficient of determination (R^2) and root mean square error (RMSE). Spectral indices such as RVI and DVI were shown to be the most reliable indices for estimating chlorophyll content in sugarcane fields, with coefficients of determination (R^2) of 0.94 and 0.93, respectively. XGB model shows the highest validation score (R^2) and lowest RMSE in both methods of BFS (0.96 and 0.14) and AFS (0.98 and 0.78), respectively. However, KNN and SVR algorithms show the lowest validation accuracy than other models. According to the results, the AFS validation score is higher than BFS in MLR, SVR, XGB and KNN. Even though, validation score of the ANN model is decreased in AFS. The findings demonstrated that the use of multispectral UAV could be utilized to estimate chlorophyll content and measure crop health status over a larger sugarcane field. This methodology will aid in real-time crop nutrition management in sugarcane plantations by reducing the need for conventional measurement of sugarcane chlorophyll content.

Keywords: chlorophyll; machine learning; multispectral imagery; remote sensing; sugarcane; UAV; vegetation indices

1. Introduction

The use of UAVs for agriculture and plant biosecurity is rapidly increasing [1–6] and the use of UAV remote sensing for precision agriculture (PA) has grown dramatically [7]. The use of unmanned aerial vehicles (UAVs) for remote sensing (RS) has developed rapidly as a method of capturing high-resolution images from the near surface of the Earth [8–13]. Several remote sensing applications have proven to be a valuable source of reflectance data for estimating various crop canopy variables relating to biophysical, physiological, or biochemical properties [14]. Many criteria of crop monitoring have already been proved to be relevant to remote sensing data and methodologies [15]. Remote sensing methods enable monitoring the agriculture field by detecting variations in the chlorophyll content with a large area for a short time [16]. Remote sensing of plant spectral responses has been demonstrated to be a promising method for capturing changes in vegetation attributes while also providing a non-destructive approach [17].

The evolving UAV platforms provide several benefits (for example, they are economical, versatile, and less affected by environmental variables), as well as their capability to collect high temporal and spatial resolution data [18]. Advances in low-altitude remote sensing technology, such as UAVs, provide a high temporal and spatial resolution solution for non-destructive, quick, and accurate assessment of various crops' biophysical parameters [19]. Satellite and manned aircraft-based remote sensing platforms can also monitor crop status across large fields and measure numerous crop and environmental parameters in real time based on precise spectral information. However, limited spatial and temporal resolution and expensive equipment costs are the major constraints in satellite remote sensing over UAV applications. Even though are highly sensitive to environmental factors [20] and UAV payload limitations and flying time is lower than those possible with satellite or manned aircraft remote sensing [20], UAV is however a viable technology and a good aerial platform for farmers, with high spatial and temporal resolution benefits [21,22].

UAV remote sensing platforms equipped with different sensors including RGB, multispectral and hyperspectral cameras have emerged as a viable option for rapid high-throughput phenotyping due to their flexibility and convenience, on-demand data access, and high spatial resolution [23,24] illustrated the application and suitability of different UAV cameras used in smart farming. RGB cameras are highly suited for the determination of canopy height and lodging. In contrast, multispectral cameras are highly suited for drought stress detection, pathogen detection, estimation of nutrients, determination of growth vigour, yield prediction, and hyperspectral and multispectral cameras are more suitable for the identification of pests and disease, weed detection and estimation of nutrient status [25].

Sugarcane (*Saccharum officinarum*) is a perennial crop commonly planted throughout the tropical and subtropical regions in the world [26]. Many natural and manmade disturbances and stressors directly impact chlorophyll content, which is the principal pigment that drives photosynthesis [16]. The accurate measurement of leaf chlorophyll concentration is critical for examining overall plant health, regulating fertilizer application, and other inputs [17]. The chlorophyll content is associated with nitrogen concentration in vegetation and is an indicator of photosynthetic activity [15]. Traditional methods for pigmentation analysis, such as spectrophotometers, leaf destruction, or high-performance liquid chromatography (HPLC), cannot quantify changes in pigmentation over a short time [19]. Furthermore, these technologies are time-consuming, costly, and assessing the health of the crops is unfeasible. As a result, these methods have limitations in monitoring crop nutritional status over a large area and greatly discourage monitoring crops. As a result, reliable, efficient, and practical methods for estimating this biophysical parameter are required [19,27,28].

1.1. Application of Remote Sensing on Sugarcane Crops

Studying the spectrometric response of leaves is essential because spectral properties are linked to non-destructive plant growth and health monitoring [29]. Plant nutritional

analysis has been made easier because spectral vegetation indices were collected from UAV images and machine learning techniques [30]. Bei Cui et al. [31] designed a new method for estimating chlorophyll content in winter wheat based on crop canopy reflectance with low sensitivity to the leaf area index (LAI) and consistent sensitivity to various crop growth situations. Gitelson et al. [32] established the conceptual model to estimate chlorophyll in maize and soybean canopies remotely, critical for regional and global carbon balance and fertilizer management. Ballester et al. [33] performed experiments to verify that the green, red vegetation index (GRVI), and the red edge ratio (RE/R) derived from UAS imagery could be used to monitor the effects of soil water status in cotton crop.

Nitrogen deficiency reduces the leaf chlorophyll concentration, enhancing the leaf's transmittance at visible wavelengths. As a result, reflected radiation from crop has been used to measure chlorophyll content. These pigments have diverse spectrum behaviour with particular absorption properties at different wavelengths, allowing the estimation of the chlorophyll concentration by remote sensing tools [19]. Spectral indices are a strong remote sensing feature for measuring plant nitrogen concentrations [34]. Therefore, spectral vegetation analysis is a viable alternative for estimating plant health [30]. On the other hand, those indices result from a combination of responses to changes in a variety of vegetation and environmental variables, such as the LAI and leaf chlorophyll content [15]. Syed Haleem Shah et al. [17] used standard statistical methodologies in conjunction with random forest regression algorithm machine learning (ML) techniques with 45 existing vegetation indices to assess the potential of hyperspectral data to estimate chlorophyll in wheat. A typical univariate regression ML analysis was used as a baseline to model the association between observed chlorophyll and the selected vegetation indices.

1.2. Machine Learning for Crop Health and Chlorophyll Content

ML algorithms have recently been applied to a variety of remote sensing applications to monitor and measure crop health and parameters [35]. ML approaches try to create an empirical relationship between the independent factors and yield, giving them the advantage of forecasting production without relying on specific crop attributes [36]. neural networks (ANN), random forests (RF), support vector machines (SVM), decision trees (DT), and other algorithms are useful for UAV-based image processing [30]. Multiple linear regression (MLR) is a statistical technique that predicts the outcome of a dependent variable using multiple independent variables. The RF Regression is a supervised learning approach for regression that uses the ensemble learning technique for remote sensing-based agricultural research projects [17,37]. The decision tree (DT) uses a tree topology to generate regression or classification models. It incrementally cuts down a dataset into smaller sections while simultaneously developing an associated decision tree. Support vector regression (SVR) is based on the same premise as SVM, but it is used to solve regression problems. Support vector regression (SVR) is a regression technique that is an extension of support vector machine (SVM). SVR develops an ideal separating hyperplane in order to distinguish classes that overlap and are not linearly separable. In this scenario, a huge modified feature space is produced to map the data and then separated along a linear boundary using kernel functions [38].

ML technologies have been used to predict crop parameters [37]. For example, winter wheat biomass estimation was carried out using the visualization approach for SVR and the investigation of influential textures [39]. Extreme gradient boosting (XGB) is a class of ensemble machine learning techniques for classification and regression predictive modelling tasks. It is an effective gradient boosting implementation that may be used for predictive regression modelling. K-Nearest neighbors (KNN) regression, RF, SVR are a non-parametric technique that approximate the connection between independent variables and the continuous outcome. Computations and mathematics are used in the ANN model to imitate human-brain processes. That of a biological nervous system influences the architecture format of the ANN models. The ANN models are composed of a complex and

nonlinear network of neurons like the real brain. Han et al. [40] showed that ANN is more effective than random forest regression in calculating maize above-ground biomass.

Zhang et al. [18] used four structure VIs and two chlorophyll VIs, as well as three regression algorithms (MLR, ANN and RF) in a maize field in Inner Mongolia, China, to measure the maize chlorophyll and vegetation indices (VIs) to crop water stress. Moran et al. [41] looked at using narrowband vegetation indices and multivariate approaches like PLSR and RF regression to estimate forest canopy chlorophyll content from airborne hyperspectral data. [26] used spectral data from a spaceborne hyperspectral image to determine sugarcane canopy nitrogen concentration spatial variation using MLR and SVR. Xu et al. [42] used six regression algorithms, including MLR, SMR, KRLS, GLM, GBM and RF, to assess the yield using a ML with UAV-lidar data in sugarcane crops. The Table 1 illustrate the different ML algorithms and their equations and optimal hyperparameter values.

Table 1. Different ML models, general equation and Hyperparameter values.

No	Models	Equation	Hyper Parameter and Optimum Values
1	MLR		normalize = True
2	RF	$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$ where; R ² : R-squared	Bootstrap = True max_depth = 10 n_estimators = 150 random_state = 11
3	DT	SS _{res} : residual sum of squares SS _{tot} : total sum of squares.	criterion = mse random_state = 0 max_depth = 9
4	SVR	$RMSE = \sqrt{\frac{\sum_{i=1}^n ((\hat{y}_i - y_i)^2)}{n}}$ where;	kernel = rbf
5	XGB	RMSE: root mean square error $\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n$ are predict values y_1, y_2, \dots, y_n are observed values n is the number of observations	learning_rate = 0.001 max_depth = 20 n_estimators = 200 use_label_encoder = False
6	KNN		n_neighbors = 23 weights = uniform algorithm = auto leaf_size = 30
7	ANN		learning_rate = 0.0001 epochs = 7000 batch_size = 100 validation

Canata et al. [43] used Sentinel-2 multi-temporal imagery data to experiment and found that the RF regression method enabled the development of predictive yield models for commercial sugarcane fields, with the authors concluding that the RF regression method was more accurate (lower RMSE and higher R²) than the MLR. RF has lately gained popularity in remote sensing research for classification and regression. The RF algorithm's variable importance plot is particularly good at detecting the most important input variables in the model [41]. Using spectral vegetation indices computed from UAV-imagery and the RF method Osco et al. [30] provided a new framework to infer nitrogen content in citrus trees at a canopy level. Feng et al. [36] used UAV-based hyperspectral imaging and ensemble learning to forecast alfalfa yield using SVR, KNN, and RF. Lee et al. [44] used three empirical methods (linear regression, RF, and SVR) to statistically connect spectral data and nitrogen levels in two corn fields in Canada. Combining machine learning techniques with spectral vegetation indices is a relatively new advanced practice to overcome the limitation of the conventional method of determining the amount of chlorophyll in plants which is time-consuming and labour-intensive [30].

Only a few studies have found a link between leaf nitrogen levels and chlorophyll content. Therefore, a major goal of this paper was to assess high-resolution multispectral

UAV images for non-destructive measurement of the chlorophyll content of sugarcane crops. There were three sub-objectives; (1) to correlate the vegetation indices with the variations of the chlorophyll in the field; (2) to compare the validation performance of the before feature selection (BFS) and after feature selection (AFS) approaches in selected ML models; and (3) to assess the prediction performance on prediction of chlorophyll content by different ML methods.

2. Methodology

2.1. Study Site

The study was carried out during September 2021 sugarcane growing season in a 1512 m² field located in Sugarcane Research Institute (SRI), Uda Walawe, Sri Lanka as shown in the Figure 1. The sugarcane variety of SL 96 128 was planted on the reddish-brown earth (RBE) in the sugarcane field.

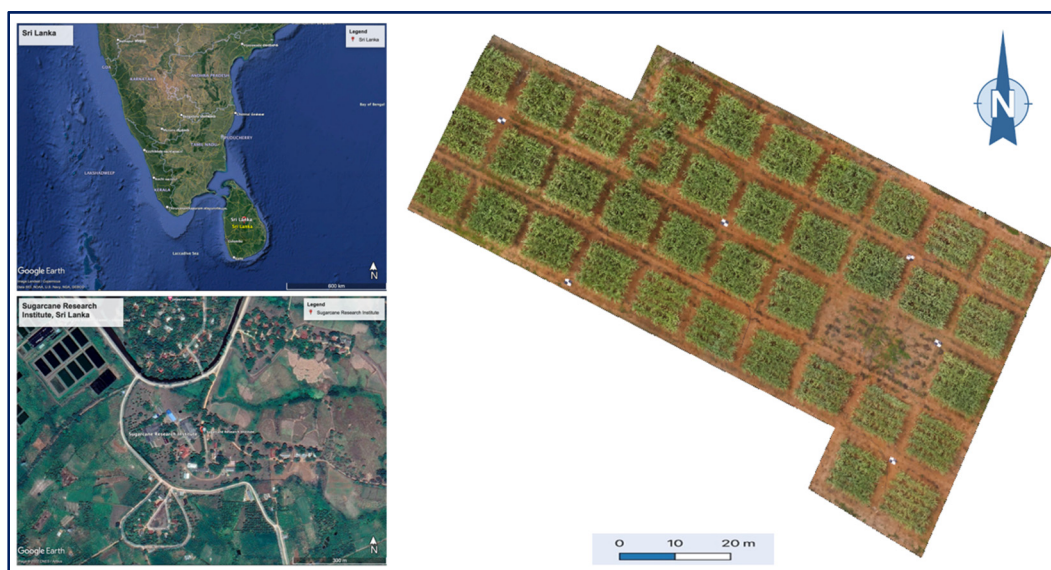


Figure 1. UAV Experimental Area at SRI.

Average climatic (Table 2) during the study period was collected from a weather station located at the SRI, Uda Walawe.

Table 2. Average climate data at the study site for September.

Parameter	Reading
Average air temperature	27.98 °C
Maximum air temperature	33.5 °C
Minimum air temperature	24.0 °C
Minimum relative humidity	50%
Maximum relative humidity	71%
Average wind speed	4.2 km/h

2.2. Experimental Design

The whole sugarcane field was allocated into twelve (12) fertilizer treatments with three replications, as shown in Figure 2 and Table 3. Altogether thirty-six (36) blocks (7 × 6 m²) were designed for each treatment, and 90 three-budded setts were planted per block. Two sampling sites of subblock (1.5 × 1.5 m²) were selected randomly in each block as average canopy area of each plant is 1.5 m².



Figure 2. Experimental design.

Table 3. Different fertilizer applications.

Treatment No	Treatment
T1	Fertilizer mixture (FM) 1 (Provide 80% from SRI recommended nutrient level)
T2	FM 1 (Provide 65% from SRI recommended nutrient level)
T3	FM 1 (Provide 50% from SRI recommended nutrient level)
T4	FM 2 (Provide 80% from SRI recommended nutrient level)
T5	FM 2 (Provide 65% from SRI recommended nutrient level)
T6	FM 2 (Provide 50% from SRI recommended nutrient level)
T7	100% from SRI recommendation
T8	80% from SRI recommendation
T9	75% from SRI recommendation + 25% Compound fertiliser
T10	65% from SRI recommendation
T11	50% from SRI recommendation
T12	Zero Fertilizer

FM 1: carbamide + triple superphosphate (TSP) + muriate of potash (MOP) + compost (1:4). FM 2: carbamide + MOP + compound fertiliser + compost (1:4).

2.3. Ground Truth Data Collection

This study used ground measurements of chlorophyll SPAD reading as references for sugarcane nutrient status, as shown in Figure 3. Chlorophyll content was collected using the SPAD-502 plus chlorophyll meter (accuracy of ± 1.0 , Konico Minolta optics Inc Osaka, Japan). Three upper side of the leaves were selected to measure the SPAD readings within the each subblock. A total of 216 SPAD readings were collected during the vegetative stage of the 5-month-old plant to build the different ML models, and the sample locations were geo-located using a Triton 2000 handheld GPS receiver (Magellan, CA, USA).

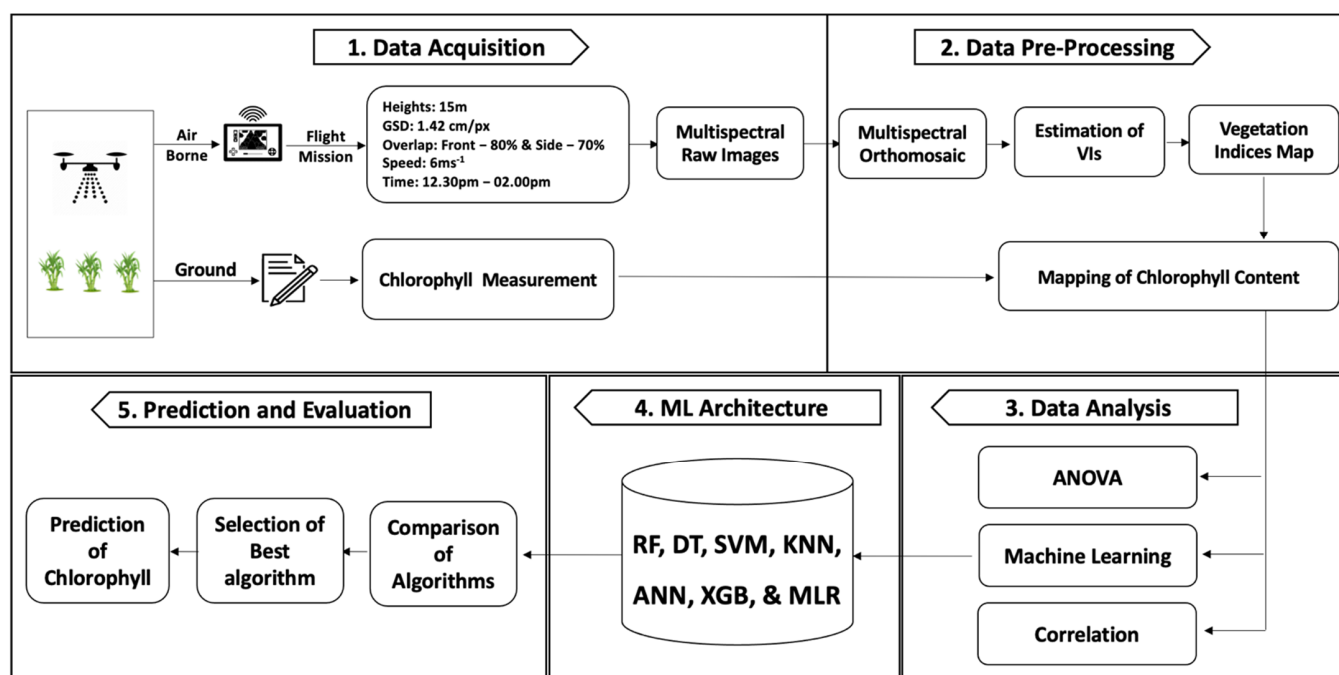


Figure 3. Primary pipeline for prediction of chlorophyll content in the sugarcane field.

2.4. Acquisition and Preprocessing of UAV Multispectral Images

A DJI P4 multispectral system (Da-Jiang Innovations (DJI), Shenzhen, Guangdong, China) was used to conduct a UAV flight mission on a sunny day between 12.30 pm and 02.00 pm (Sri Lankan standard time) during the sugarcane growing season. The visible to the near-infrared spectral range of the DJI P4 multispectral camera has five bands (blue, green, red, red edge, and near infrared) at 450.0, 560.0, 650.0, 730.0 and 840.0 nm, respectively. The flight altitude above ground, speed, and ground sample distance, were 15 m, $6 \text{ m} \cdot \text{s}^{-1}$ and 1.42 cm, respectively. The front and side overlap of images on the flight line was 80 % and 70 %, respectively, as shown in Table 4. Six ground control points (GCPs) were used to improve geolocation accuracy for post-image processing. The image mosaic processing was carried out with Agisoft Metashape (Version-1.6.6, Agisoft LLC, Petersburg, Russia).

Table 4. UAV Flight Mission.

Height	15m
Ground Sampling Distance (GSD)	1.42 cm/px
Speed	6 ms^{-1}
Overlap	Front-80% & Side-70%
Time	12.30–02.00 p.m.

2.5. Estimation of Vegetation Indices

The values of reflectance in the red, green, blue, red edge and NIR portions of the electromagnetic spectrum of UAVs were used to generate several VIs. Twenty-four (24) VIs were estimated, as shown in Table 5 to demonstrate the feasibility of calculating sugarcane vegetation indices to predict the chlorophyll content. Two rectangle regions of interest (ROI) were identified based on the GPS coordinates collected during ground truth measurement in each block on the aerial vegetation index map. The average VIs inside the ROI was determined. Generation of vegetation Indices and extraction of index values were performed using the Open-Source Geographic Information System of QGIS (version-3.20) [45].

Table 5. Vegetation indices calculation.

No	Vegetation Indices	Formula	Purpose	References
01	Normalized Difference Vegetation Index (NDVI)	$NDVI = \frac{NIR - R}{NIR + R}$	Estimation of vegetation biomass,	[43]
02	Green Normalized Difference Vegetation Index (GNDVI)	$GNDVI = \frac{NIR - G}{NIR + G}$	Estimation of vegetation fraction, Estimation of productivity and assess phylogenetic heritability, Discrimination of field-grown olive cultivars	[44,46]
03	Normalized Difference Red Edge Index (NDRE)	$NDRE = \frac{NIR - Red\ Edge}{NIR + Red\ Edge}$	Detect Differences in Vegetation and Chlorophyll Content., Early stress detection, Growth monitoring	[47–49]
04	Leaf Chlorophyll Index (LCI)	$LCI = \frac{NIR - Red\ Edge}{NIR + R}$	Early detection of pine wilt disease	[50]
05	Difference Vegetation Index (DVI)	$DVI = NIR - R$	Prediction of green LAI of crop canopies, Comparison of different reflectance indices for vegetation analysis, discriminate field-grown olive cultivars	[46,51,52]
06	Ratio Vegetation Index (RVI)	$RVI = \frac{NIR}{R}$	Early detection of pine wilt disease	[50]
07	Enhanced Vegetation Index (EVI)	$EVI = \frac{2.5(NIR - R)}{NIR + 6R - 7.5B + 1}$	Early detection of pine wilt disease	[50]
08	Triangular Vegetation Index (TVI)	$TVI = 60(NIR - G) - 100(R - G)$	Early detection of pine wilt disease	[50]
09	Green Chlorophyll Index (GCI)	$GCI = \frac{NIR}{G} - 1$	Estimation of leaf area index and green leaf biomass	[52,53]
10	Green Difference Vegetation Index (GDVI)	$GDVI = NIR - G$	Generalized Difference Vegetation Index (GDVI) estimate productivity and assess phylogenetic heritability	[45,46,54]
11	Normalized Green Red Difference Index (NGRDI)	$NGRDI = \frac{(G - R)}{(G + R)}$	Monitoring of crop biomass	[46,55]
12	Modified Soil-Adjusted Vegetation Index (MSAVI)	$MASVI = \frac{(2 \times NIR + 1 - \sqrt{(2 \times NIR + 1)^2 - 8 \times (NIR - R)})}{2}$	Significant remote sensing vegetation indices	[56]
13	Atmospherically Resistant Vegetation Index (ARVI)	$ARVI = \frac{(NIR - (R - 2 \times (B - R)))}{(NIR + (R - 2 \times (B - R)))}$	Significant remote sensing vegetation indices	[56]
14	Structure Insensitive Pigment Index (SIPI)	$SIPI = \frac{(NIR - B)}{(NIR - R)}$	Mapping the sugarcane	[57]
15	Optimized Soil-Adjusted Vegetation Index (OSAVI)	$OSAVI = \frac{1.16 \times (NIR - R)}{NIR + R + 0.16}$	Extraction of land cover information	[58]
16	Green Optimized Soil Adjusted Vegetation Index (GOSAVI)	$GOSAVI = \frac{NIR - G}{NIR + G + 0.16}$	Extraction of land cover information	[58]
17	Excess Green (ExG)	$ExG = \frac{2G - R - B}{R + G + B}$	Vine diseases detection	[59]
18	Excess Red (ExR)	$ExR = \frac{1.4R - G}{R + G + B}$	Vine diseases detection	[59]
19	Excess Green Red (ExGR)	$ExGR = ExG - ExR$	Vine diseases detection	[59]
20	Green Red Vegetation Index (GRVI)	$GRVI = \frac{R - G}{R + G}$	Vine diseases detection	[59]
21	Normalized Difference Index (NDI)	$NDI = \frac{G - R}{G + R}$	Vine diseases detection	[59]
22	Red Green Index (RGI)	$RGI = \frac{R}{G}$	Vine diseases detection	[59]
23	Enhanced Normalized Difference Vegetation Index (ENDVI)	$ENDVI = \frac{((NIR + G) - (2B))}{((NIR + G) + (2B))}$	Mapping the sugarcane	[57]
24	Simple Ratio Index (SRI)	$SRI = NIR/R$	Evaluation of the saturation property of vegetation indices	[60,61]

2.6. Machine Learning Modelling and Statistical Analysis

Statistical analysis was used to examine and establish an association between the UAV-derived vegetation indices and ground-truthing SPAD reading through different machine learning modelling by using Python (version 3.8.10). One of the feature selection

techniques of Pearson's correlation coefficient was utilized to select which vegetation indices were most sensitive to chlorophyll, and the highest correlation coefficients (R^2) values from reflectance features were used to develop machine learning algorithms to predict the sugarcane chlorophyll content accurately. In this study, seven (07) machine learning regression algorithms, MLR, RF, DT, SVR, XGB, KNN and ANN were compared to predict the sugarcane chlorophyll concentration based on VIs derived from reflection images. The root means square error (RMSE) and the coefficient of determination or validation score (R^2) were calculated for training and validation to compare and select the best-fit algorithm for chlorophyll prediction in the sugarcane field. In statistics and machine learning, feature selection refers to the process of selecting a subset of relevant features (predictors and variables) for inclusion in a model. It is the process of automatically selecting the data qualities (such as columns in tabular data) that are most significant and pertinent to the predictive modelling challenge at hand. It is mean that to minimize the number of input variables to those that are deemed to be most beneficial in predicting the target variable. According to the previous studies mentioned in the Table 4, 24 VIs were selected to correlate the ground truth measurement. After estimating the correlation values for all VIs (before feature selection), 15 VIs were selected (after feature selection) based on the Pearson correlation values greater than 50% (± 0.5) to improve the model performance and reduce the training time for the development of ML models. Finally total of 216 samples were used to build the different ML models.

3. Results

3.1. One Way ANOVA Statistical Analysis for Different Treatments and Chlorophyll Content

A one-way ANOVA test was performed to estimate the significant relationship between all twelve (12) fertilizer treatments and sugarcane chlorophyll content. The result shows a significant interaction ($p = 0.001$) between all fertilizer treatment and chlorophyll content. Figure 4 shows the comparing the treatment means and variability of chlorophyll reading, and Figure 5 shows the quantile-quantile (Q-Q) plot confirming that the data were adequately close to the theoretical reference line, representing a soundly model fit.

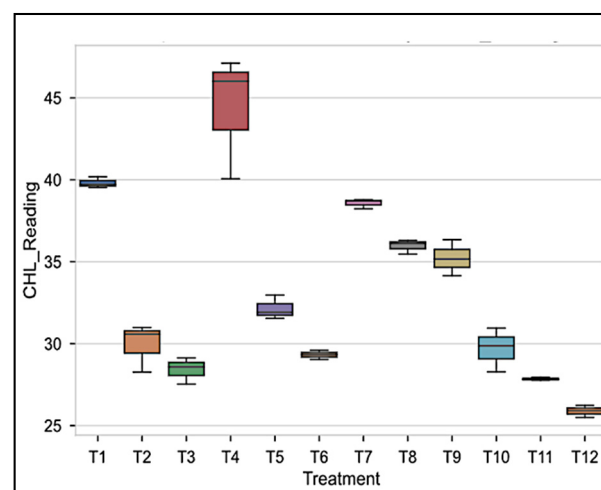


Figure 4. Comparison of the box plots for one way ANOVA.

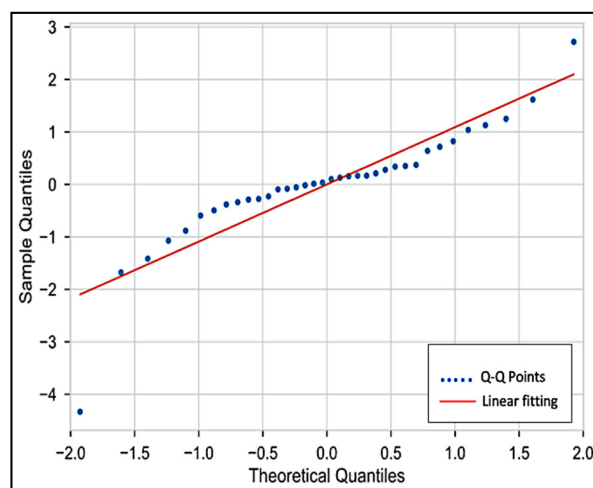


Figure 5. Normal Q-Q plot for the observed sample against theoretical quantiles.

3.2. Correlation between Vegetation Indices and Sugarcane Chlorophyll Content

The Pearson's correlation coefficients (R^2) for the relationship between VIs and sugarcane chlorophyll content are shown in Figure 6 and detailed correlation matrix shown in Figure A1 (Appendix A). Pearson's correlation test was performed to select the essential features crucial for training the ML algorithms. The RVI showed the highest positive correlation with the chlorophyll content (R^2 value: 0.94), and DVI also had stronger correlations with chlorophyll content (R^2 value: 0.93). Next to the RVI and DVI, other vegetation indices such as NDRE, GNDVI, LCI, EVI, and NDVI showed positive correlation coefficients of 0.86, 0.86, 0.85, 0.84, and 0.82, respectively.

3.3. Prediction of Sugarcane Chlorophyll Content by Using Machine Learning Algorithms

Different ML techniques including MLR, RF, DT, SVR, XGB, KNN and ANN were developed in two methods. The first is before feature selection (BFS)—training the algorithms with all twenty-four (24) vegetation indices and five (05) spectral bands; The second method is after feature selection (AFS)—training the algorithms with selected fifteen (15) vegetation indices. The two methods were compared with the estimated coefficient of determination (R^2) and root mean square error (RMSE), as shown in Figure 7.

As shown in Table 6, XGB model shows the highest validation score (R^2) and lowest RMSE in both methods of BFS (0.96 and 0.14) and AFS (0.98 and 0.78), respectively. As for RF, both R^2 values derived from the validation data set were lower than the XGB. Also, the MLR model also shows a good training and validation score in both methods. However, KNN and SVR algorithms show the lowest validation accuracy than other models. When comparing the two approaches, the AFS validation score increases in MLR, SVR, XGB and KNN. Even though RF and DT show no changes in validation score in both methods, the validation score of the ANN model decreases in AFS.

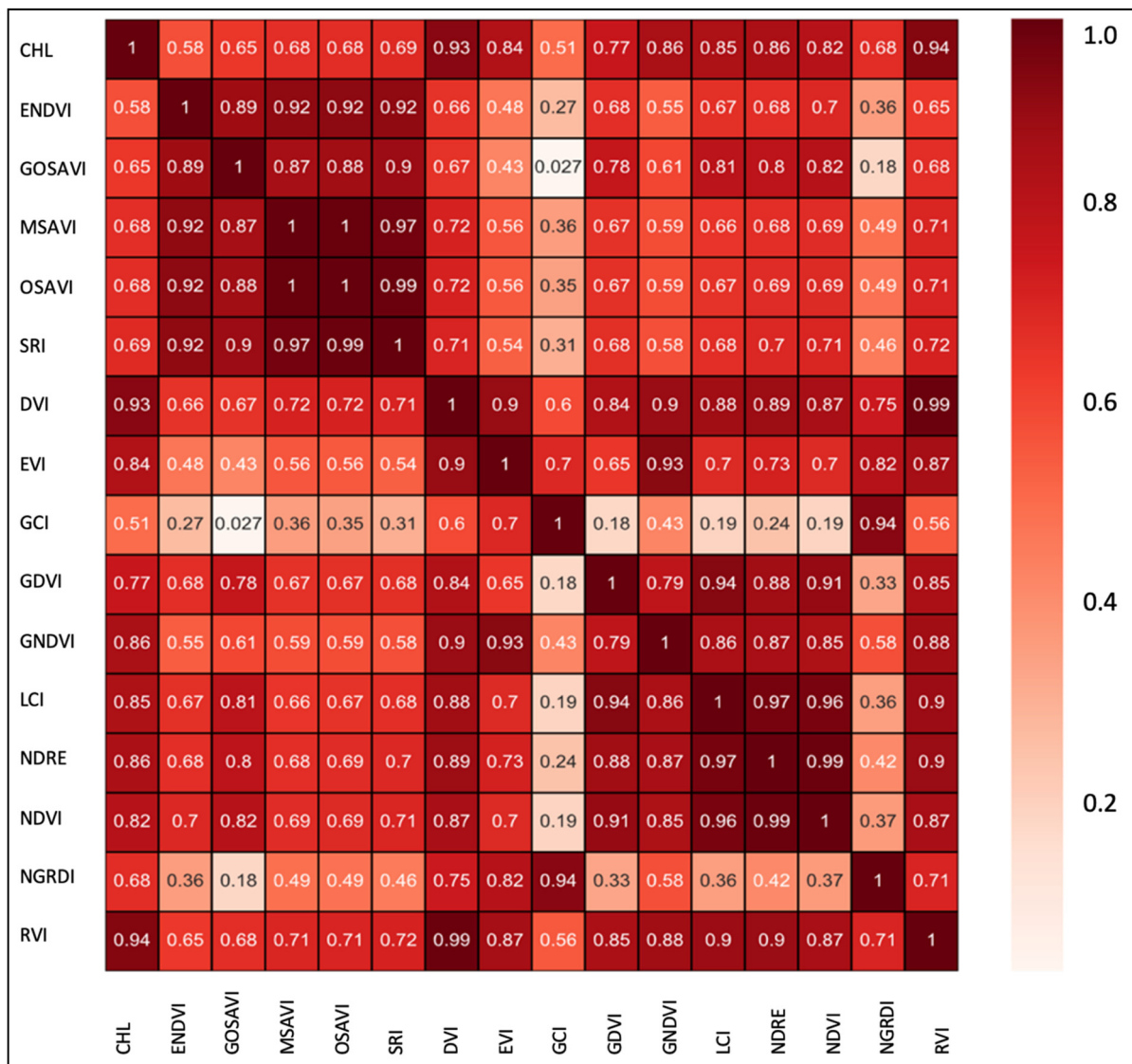


Figure 6. Correlation matrix plot between chlorophyll content and different vegetation indices for the AFS method.

Table 6. Performance comparison of the different ML models.

ML Model	R ² (Training)		R ² (Validation)		RMSE	
	BFS	AFS	BFS	AFS	BFS	AFS
MLR	0.95	0.93	0.82	0.83	2.14	2.09
RF	0.99	0.99	0.95	0.95	1.11	1.09
DT	1.00	1.00	0.94	0.94	1.21	1.25
SVR	0.74	0.85	0.65	0.78	3.02	2.37
XGB	0.99	0.99	0.96	0.98	0.14	0.78
KNN	0.69	0.73	0.68	0.75	3.42	3.18
ANN	0.98	0.98	0.87	0.76	1.83	2.48

MLR: Multiple linear regression; RF: Random Forest regression; DT: Decision tree regression; SVR: Support vector regression; XGB: eXtreme Gradient Boosting; KNN: k-nearest neighbors; ANN: Artificial neural network; BFS: Before feature selection; AFS: After feature selection; R²: Coefficient of determination; RMSE: Root mean square error.

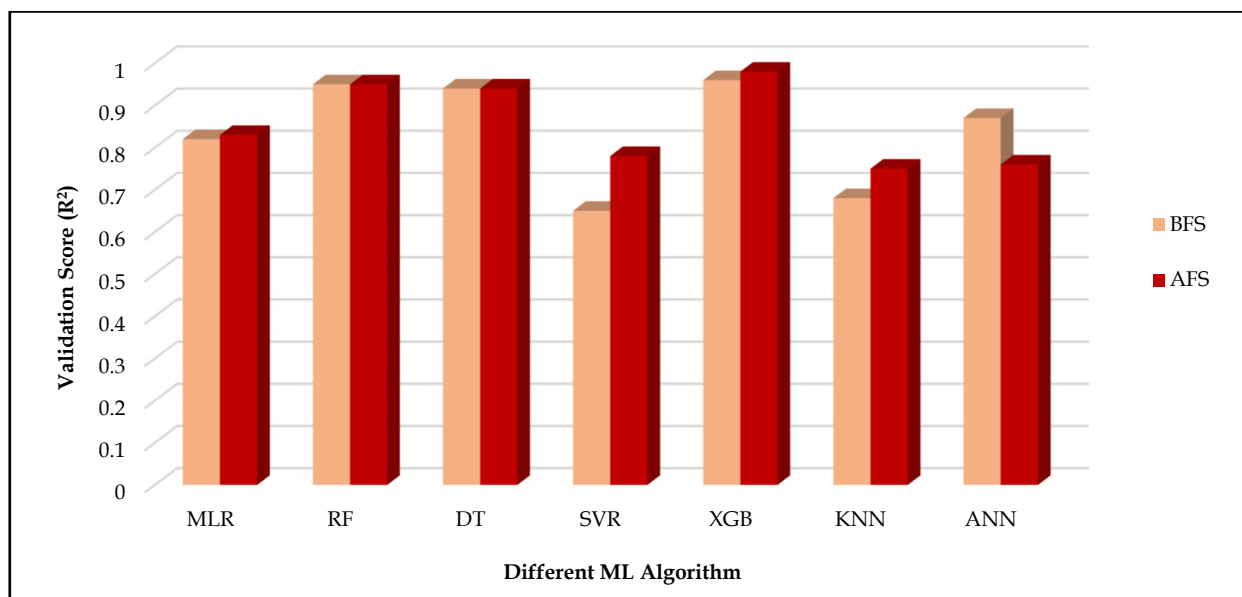


Figure 7. Comparison of validation (R^2) score for all ML models.

4. Discussion

In this study, we compared canopy level multispectral data to estimate leaf chlorophyll in sugarcane crops using different ML architecture. Chlorophyll has long been thought to be the most crucial pigment for detecting nutritional stress. When sugarcane canopy structure was more responsive to nutrient stress, the results demonstrated that the chlorophyll content could only assess sugarcane nutrient stress. Chlorophyll concentration drops when nutrient stress occurs and causes structural or colour changes identified as visual nutrient stress symptoms. The use of ML architectures including MLR, RF, DT, SVR, XGB, KNN, and ANN comparing (1) all spectral bands and (2) selected VIs as well as regressions analysis using existing VIs were investigated.

4.1. Basic Statistical Analysis

An initial ANOVA test was used to determine whether or not the treatment results of an experiment are significant, and F-distribution was used to compare two means from two independent variables of VIs using a one-way ANOVA [62]. As shown in Figure 4, the result is statistically significant, which indicates that the two means are unequal. This test confirms the conducted fertilizer treatments are significantly different from each other. Therefore, we can confirm that all the fertilizer treatments show significant variation among them, which is important to develop efficient ML models to predict sugarcane chlorophyll content. After confirming the ANOVA outputs, the pearson's correlation coefficients (R^2) for the relationship between UAV-derived VIs and sugarcane chlorophyll content are estimated to select the essential features crucial for training the ML algorithms. The highly correlated input variables including RVI, DVI, NDRE, GNDVI, LCI, EVI, and NDVI are linked with the target. In this study, we used an absolute number, such as 0.5, as the variable selection threshold. If the predictor variables are found to be associated, the variable with the lowest correlation coefficient value with the target variable is discarded. However, other features selection techniques such as chi-square test, Fisher's score, variance threshold, mean absolute difference (MAD), forward feature selection, and backward feature elimination can also be used in different ML studies. Therefore, future studies can be compared the different feature selection methods for the forecast of chlorophyll to find the best prototype model [63].

4.2. Machine Learning Approach Using Multispectral Bands and VIs

The ML models evaluated; MLR, RF, DT, SVR, XGB, KNN, and ANN are all good at handling a continuous dependent variable that is correlated with VIs. The use of spectral vegetation indices in conjunction with machine learning models proved to be an effective method for predicting chlorophyll content consistent with [64], and spectral indices have proven to be an essential technique for evaluating nitrogen [14]. This study did not remove the soil from the reflectance map to estimate vegetation indices because the crop completely covered the soil during the experimental period. However, it is necessary to remove the soil for estimation of VIs during the early stage of sugarcane crops as an aerial map can be shown the soil between the sugarcane crops. We used 80% of the available VIs as input training data and 20% as validation data to estimate the machine learning model's performance on new data. Best fit line plots were generated to compare all ML models using both BFS and AFS methods, as illustrated in Figure 7. The line of best fit is a line that runs through a scatter plot of data points and best reflects the relationship between them [65]. The regression ML analysis output can be used to predict the chlorophyll content over variation in VIs. The red line in Figure 8 is referred to as the best fit straight line for each model.

Figure 9 shows the learning curves of MLR, RF, DT, SVR, XGB, KNN and ANN to evaluate the model learning performance over training instances. The shadow green and blue represent the standard deviation of accuracy, while the lines show the mean accuracy values in the proposed models. Learning curves are a common diagnostic tool for algorithms that learn progressively from a training dataset in ML [66]. The model's performance improves over time, indicating that the model is learning and improving [67].

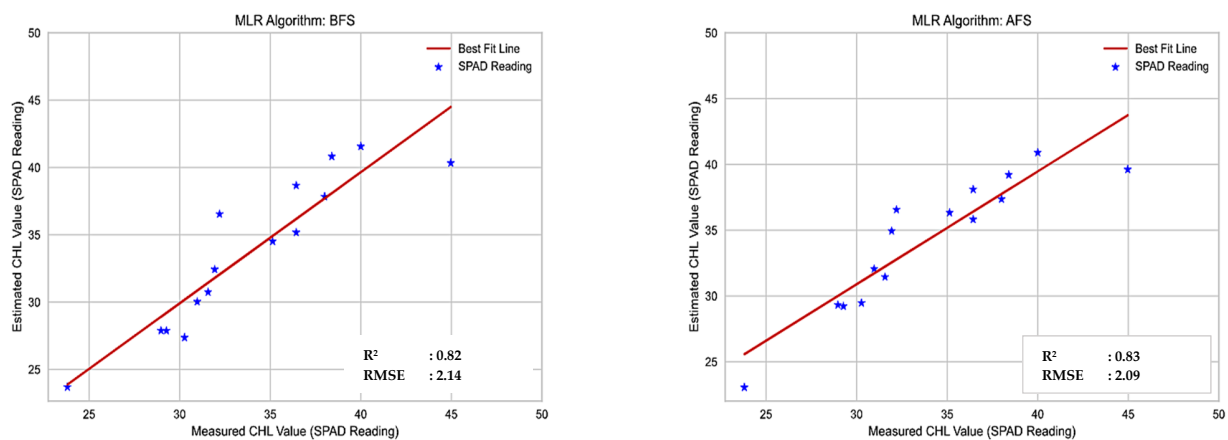


Figure 8. Cont.

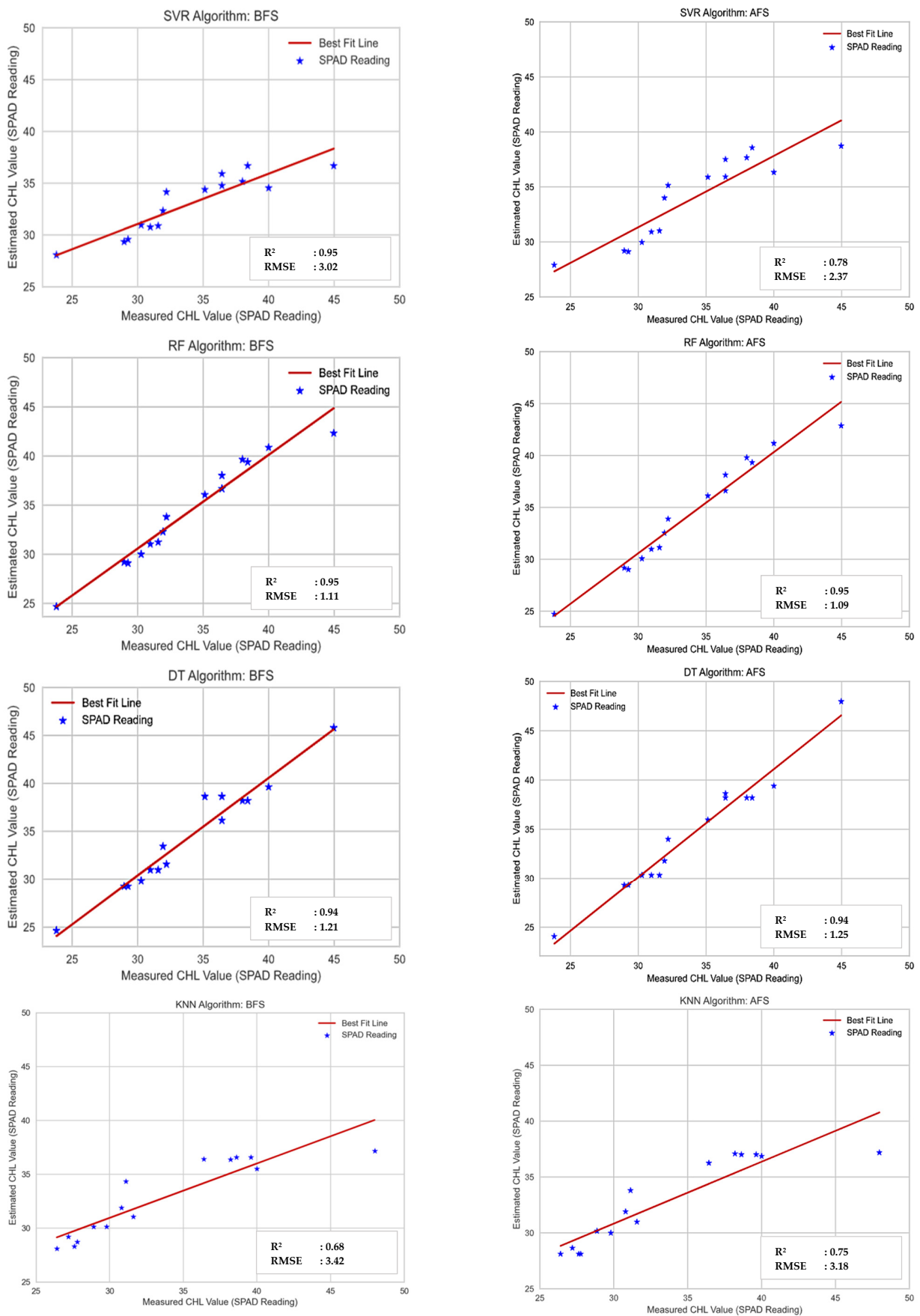


Figure 8. Cont.

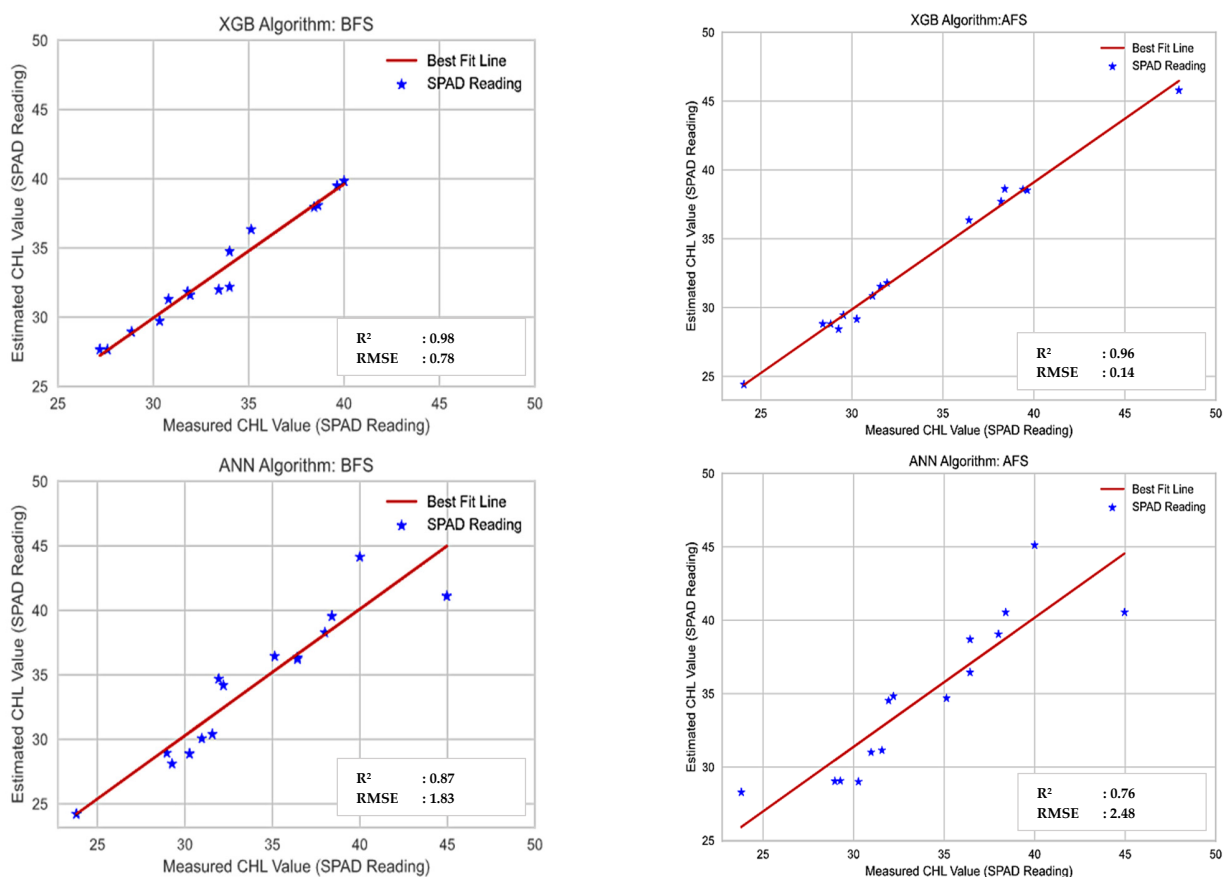


Figure 8. Best fit line of different ML models for left column BFS and right column AFS methods.

When all spectral bands were utilized as input predictors, the results showed that the XGB technique delivers a higher retrieval accuracy than other models. XGB has recently been demonstrated to be a highly effective machine learning technique for mapping in RS, and it is capable of performing well even with limited training data [68,69]. Yang et al., 2021 conducted experiments on wheat SPAD estimation utilising cluster-regression algorithms using UAV hyperspectral data, and the results indicated that the XGB model beat the random forest model somewhat in estimating wheat SPAD. Therefore, the XGB algorithm may be used for fertiliser treatments in precision agriculture [70]. Also, MLR, RF, and DT models show a good training and validation score in both methods of BFS and AFS. Although the ANN outperformed the SVM and LR algorithms, it produced results that were considered inferior to the RF and DT methods [71,72]) because ANN approaches were favored over other types of spectral information for predicting crop nitrogen stress and mapping vegetation. Additionally, different types of plant stress have been identified using ANNs and multispectral data [73]. We observed that XGB, MLR, RF, DT show almost the same training and validation score in this study. However, The RF method shows the best accuracy of 90% than other models of SVM and MLR in the previous study to estimate the chlorophyll content conducted by Osco et al. [42].

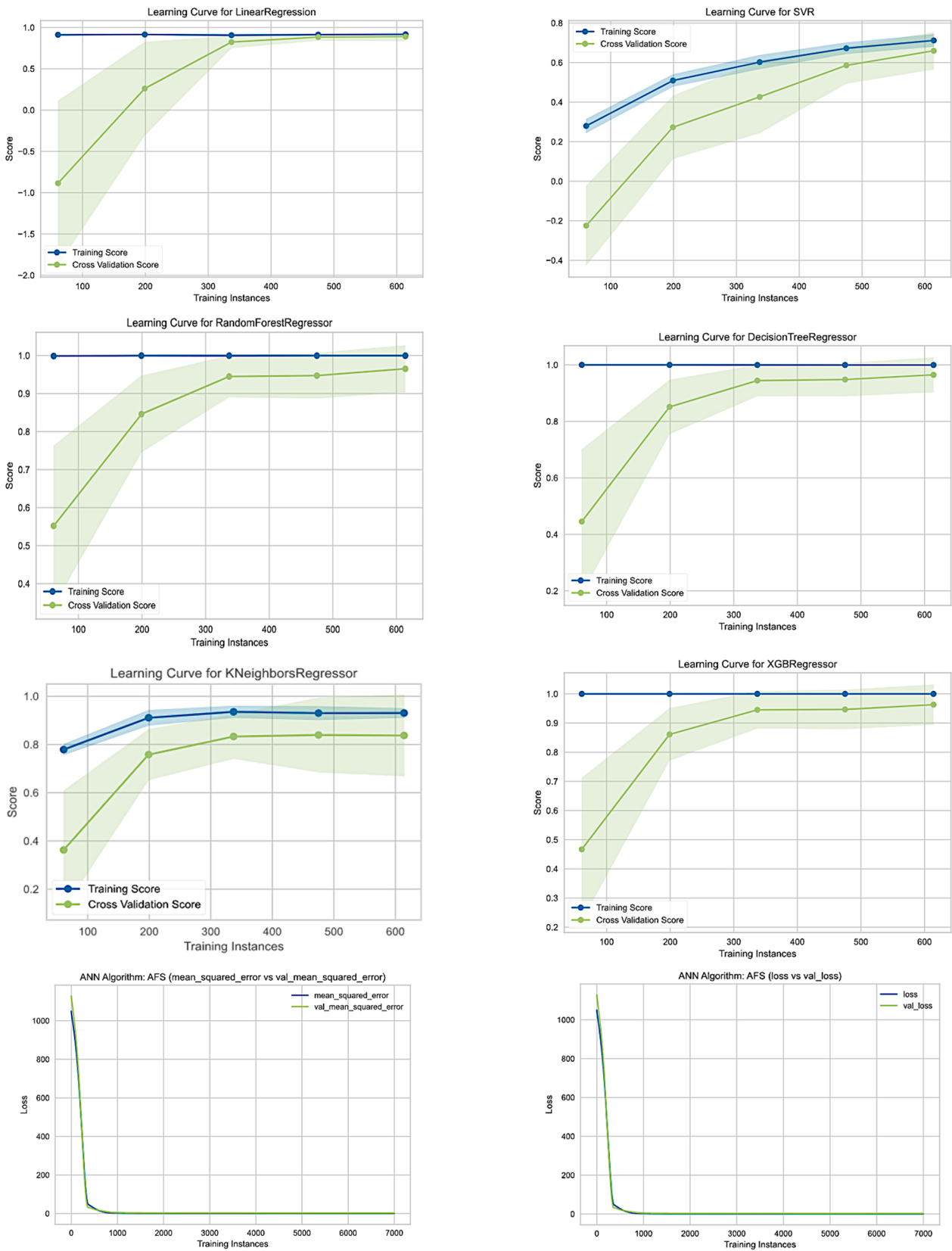


Figure 9. Learning curve of different ML models in AFS methods.

Our findings revealed that feature selection techniques can increase prediction accuracy in many models, including MLR, SVR, XGB and KNN however, these techniques were less important in the ANN model. This procedure is carried out to reduce the number

of input features while maintaining the model's predicted accuracy. We looked at a total of 24 spectral indices, which is more than prior research of this type have done [42]. In previous studies of Ballester et al. [74] and Zeng & Chen, [75], when a single VI was used to create an association with chlorophyll using basic linear regressions, the R^2 of the generated relationships had significant fluctuations in values. When using a lower number of spectral indices, the XGB model helped to increase the algorithm's accuracy [42]. This suggests that the number of spectral indices utilized might be reduced while still obtaining extremely accurate results [42]. This information is critical since it aids new research in reducing the amount of processed data, which has an impact on training and testing times [42].

SVMs and KNN are frequently utilised when scientists are confronted with a huge number of features and a high degree of sparsity [74]. Although the SVM and KNN algorithm's prediction accuracy was lower than that of the other algorithms in this investigation due to the selection of important features of this study. Previous research has also indicated an increase in SVM and KNN performance when specified variables are used [74]. This could be because picking relevant variables improves the SVM and KNN performance by increasing its interpretability, computational efficiency, and generalisation performance [74]. Furthermore, SVM and KNN algorithms are more sensitive to the quality of data than other algorithms which may lead to reduced performance of this prediction model.

This study is very important for the current fertiliser trial at SRI at Uda Walawe because chlorophyll measurement should be taken every two weeks to compare and analyse the effect of variation in different fertiliser treatments. However, it is very difficult and needs more time with labourers using SPAD meters. Therefore, this proposed ML model can be used to measure the chlorophyll content every two weeks at a large sugarcane field if we use UAV and ML techniques. Further, measuring chlorophyll by SPAD may be produced inaccurate reading due to leaf structure, water content and leaf pigment distribution [76]. Environmental factors including light intensity can also affect the light transmittance of a leaf, resulting in incorrect measurement of chlorophyll content. Therefore, the application of UAV, multispectral camera and AI can be an effective solution for fertiliser trials for sugarcane crops.

4.3. Limitations of the Experimental and Modelling Approach

The small number of samples (216 samples) is a limitation of this study. Though RF is suitable for modest amounts of sampling data, the RF model's performance is linked to the sample size, and the more sample points there are, the more accurate the forecasts [76], the predicted model may not be suitable for discriminating crop chlorophyll content from different growth stages as this research was done for 5-month-old sugarcane crops. In the modelling approach, manual hyper parameter tuning was performed to obtain the best model for all the ML algorithms including MLR, RF, DT, SVR, XGB, KNN and ANN, so future work using grid-searching, which is the process of scanning the data to align best parameters for a given model with minimal human effort, can be employed. This study focused on regression ML models. Therefore, further studies should be needed to develop the chlorophyll prediction model by ML and DL classification techniques with grid searching methods for different stages of sugarcane crops and various environmental conditions.

5. Conclusions

Chlorophyll is an important crop biophysical feature to measure crop health and create early predictions. This current study looked at the viability of using multispectral UAV images to predict the chlorophyll content of sugarcane crops in SRI, Sri Lanka. A SPAD chlorophyll meter acquired ground-truthing data of the sugarcane chlorophyll contents to correlate the different vegetation indices for ML. Different ML models were compared for several vegetation indices and the chlorophyll content of sugarcane crops to construct a prediction model for sugarcane chlorophyll content. Among the other indices utilized in the study, RVI and DVI revealed a strong and positive correlation with the chlorophyll content of sugarcane crops. The results show that the XGB technique delivered

a higher retrieval accuracy than other models when all spectral bands were utilized as input predictors. The most important finding of this study is that spectral signals derived from space multispectral data offer useful information for quantifying sugarcane chlorophyll content over greater geographic areas for implementing proper farm management. Due to practical constraints, the agronomical approach of collecting leaf tissue and performing chemical analysis in the laboratory is time consuming and spatially limited. This research and the use of UAVs with AI can positively impact fertilization procedures and lead to more accurate yield projections. With the success of predicting the chlorophyll content across larger geographic areas using spaceborne multispectral data, cane growers will be able to monitor the nutritional state of their sugarcane early and address nutrient deficient areas with appropriate management. In further work on the proposed approach for estimating chlorophyll content, this has to be tested in different sugarcane fields with different varieties and must be validated for the different growing stages of sugarcane crops.

Author Contributions: A.N. conducted the UAV flight mission, analysis and prepared the manuscript for final submission as a corresponding author. F.G. provided overall supervision and contributed to the writing and editing. A.S.A.S. provide the technical guidance to conduct the UAV flight mission and research design and provide the feedbacks on draft manuscript. U.W.L.M.K., H.A.S.W. and B.R.K. developed the experimental design in the field and carried out the fieldwork for ground sample collection. The version of the manuscript has been read and approved by all authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding and the APC was funded by Queensland University of Technology (QUT).

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors thank the Sugarcane Research Institute, Sri Lanka for giving the permission to conduct the UAV flight mission and ground truth data collection. We are indebted to the anonymous reviewers for their insightful remarks on our article.

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

Appendix A

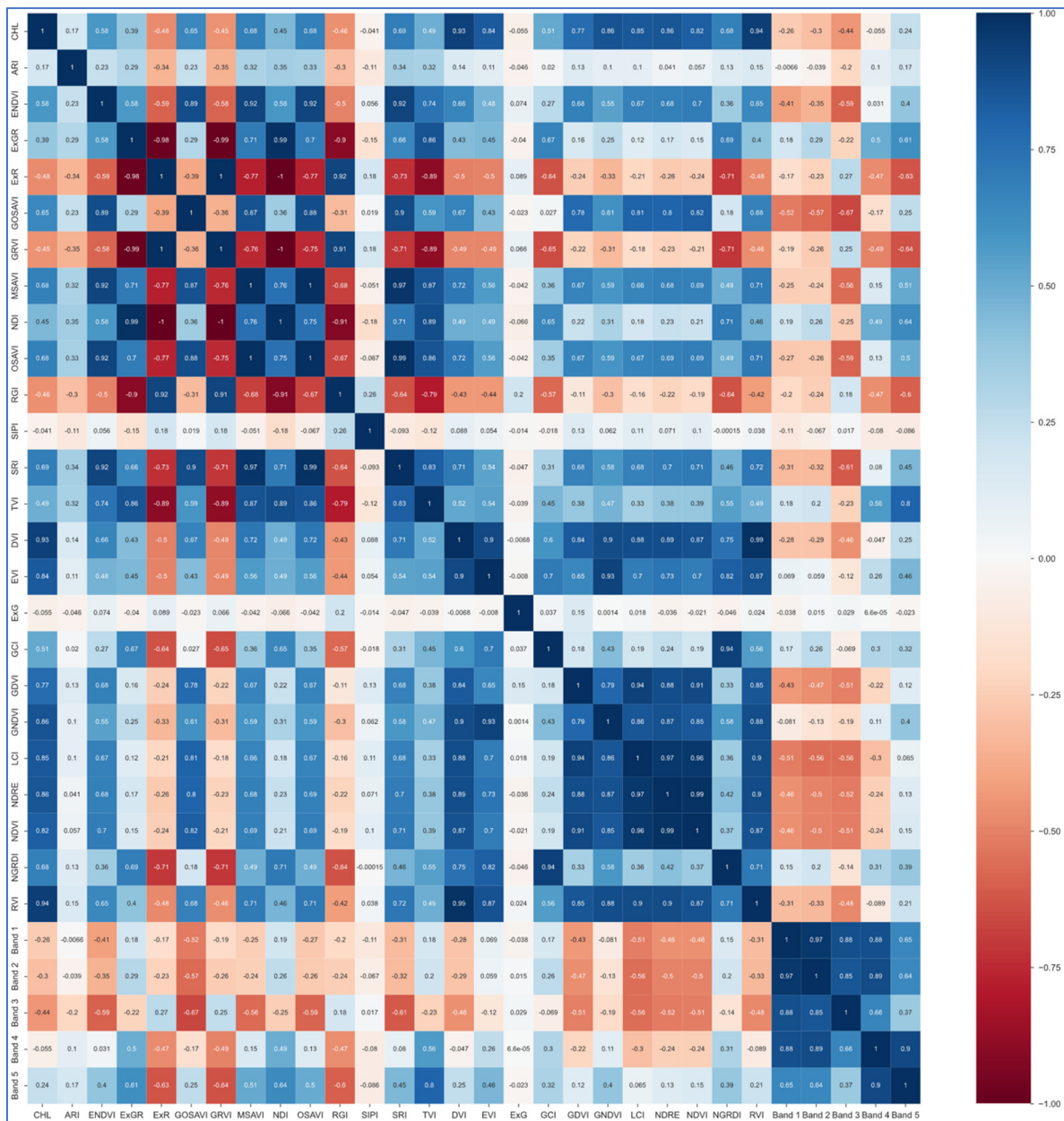


Figure A1. Correlation matrix plot between chlorophyll content and different vegetation indices.

References

1. Vanegas, F.; Bratanov, D.; Powell, K.; Weiss, J.; Gonzalez, F. A Novel Methodology for Improving Plant Pest Surveillance in Vineyards and Crops Using UAV-Based Hyperspectral and Spatial Data. *Sensors* **2018**, *18*, 260. [CrossRef] [PubMed]
2. Thomas, J.E.; Wood, T.A.; Gullino, M.L.; Ortu, G.; Thomas, J.E.; Wood, T.A. Diagnostic Tools for Plant Biosecurity. In *Practical Tools for Plant and Food Biosecurity*; Springer: Cham, Switzerland, 2017; pp. 209–226. [CrossRef]
3. Mcfadyen, A.; Gonzalez, L.F.; Campbell, D.A.; Eagling, D. *Evaluating Unmanned Aircraft Systems for Deployment in Plant Biosecurity*; Queensland University of Technology: Brisbane City, QLD, Australia, 2014. [CrossRef]
4. Puig Garcia, E.; Gonzalez, F.; Hamilton, G.; Grundy, P. Assessment of crop insect damage using unmanned aerial systems: A machine learning approach. In Proceedings of the MODSIM 2015, 21st International Congress on Modelling and Simulation, Gold Coast, Australia, 29 November–4 December 2015. Available online: <http://www.mssanz.org.au/modsim2015/F12/puig.pdf> (accessed on 14 January 2022).
5. Hu, Y.; Wilson, S.; Schwessinger, B.; Rathjen, J.P. Blurred lines: Integrating emerging technologies to advance plant biosecurity. *Curr. Opin. Plant Biol.* **2020**, *56*, 127–134. [CrossRef] [PubMed]

6. Sandino, J.; Pegg, G.; Gonzalez, F.; Smith, G. Aerial Mapping of Forests Affected by Pathogens Using UAVs, Hyperspectral Sensors, and Artificial Intelligence. *Sensors* **2018**, *18*, 944. [[CrossRef](#)]
7. Zhang, T.; Su, J.; Liu, C.; Chen, W.-H. State and parameter estimation of the AquaCrop model for winter wheat using sensitivity informed particle filter. *Comput. Electron. Agric.* **2020**, *180*, 105909. [[CrossRef](#)]
8. Seyyedhasani, H.; Digman, M.; Luck, B.D. Utility of a commercial unmanned aerial vehicle for in-field localization of biomass bales. *Comput. Electron. Agric.* **2020**, *180*, 105898. [[CrossRef](#)]
9. Nebiker, S.; Annen, A.; Scherrer, M.; Oesch, D. A light-weight multispectral sensor for micro UAV—Opportunities for very high resolution airborne remote sensing. *Int. Arch. Photogramm. Remote Sens. Spat.-Form. Sci.* **2008**, *37*, 1193–1200.
10. Yue, J.; Lei, T.; Li, C.; Zhu, J. The Application of Unmanned Aerial Vehicle Remote Sensing in Quickly Monitoring Crop Pests. *Intell. Autom. Soft Comput.* **2012**, *18*, 1043–1052. [[CrossRef](#)]
11. Aasen, H.; Honkavaara, E.; Lucieer, A.; Zarco-Tejada, P.J. Quantitative Remote Sensing at Ultra-High Resolution with UAV Spectroscopy: A Review of Sensor Technology, Measurement Procedures, and Data Correction Workflows. *Remote Sens.* **2018**, *10*, 1091. [[CrossRef](#)]
12. Casagli, N.; Frodella, W.; Morelli, S.; Tofani, V.; Ciampalini, A.; Intrieri, E.; Raspini, F.; Rossi, G.; Tanteri, L.; Lu, P. Spaceborne, UAV and ground-based remote sensing techniques for landslide mapping, monitoring and early warning. *Geoenviro. Disasters* **2017**, *4*, 9. [[CrossRef](#)]
13. Xiang, H.; Tian, L. Development of a low-cost agricultural remote sensing system based on an autonomous unmanned aerial vehicle (UAV). *Biosyst. Eng.* **2011**, *108*, 174–190. [[CrossRef](#)]
14. Hansen, P.M.; Schjoerring, J.K. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sens. Environ.* **2003**, *86*, 542–553. [[CrossRef](#)]
15. 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](#)]
16. Hoepfner, J.M.; Skidmore, A.K.; Darvishzadeh, R.; Heurich, M.; Chang, H.-C.; Gara, T.W. Mapping Canopy Chlorophyll Content in a Temperate Forest Using Airborne Hyperspectral Data. *Remote Sens.* **2020**, *12*, 3573. [[CrossRef](#)]
17. 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](#)]
18. Zhang, L.; Han, W.; Niu, Y.; Chávez, J.L.; Shao, G.; Zhang, H. Evaluating the sensitivity of water stressed maize chlorophyll and structure based on UAV derived vegetation indices. *Comput. Electron. Agric.* **2021**, *185*, 106174. [[CrossRef](#)]
19. Tahir, M.N.; Naqvi, S.Z.A.; Lan, Y.; Zhang, Y.; Wang, Y.; Afzal, M.; Cheema, M.J.M.; Amir, S. Real time estimation of chlorophyll content based on vegetation indices derived from multispectral UAV in the kinnow orchard. *Int. J. Precis. Agric. Aviat.* **2018**, *1*, 24–31. [[CrossRef](#)]
20. Paneque-Gálvez, J.; McCall, M.K.; Napoletano, B.M.; Wich, S.A.; Koh, L.P. Small Drones for Community-Based Forest Monitoring: An Assessment of Their Feasibility and Potential in Tropical Areas. *Forests* **2014**, *5*, 1481–1507. [[CrossRef](#)]
21. Jang, G.; Kim, J.; Yu, J.-K.; Kim, H.-J.; Kim, Y.; Kim, D.-W.; Kim, K.-H.; Lee, C.W.; Chung, Y.S. Review: Cost-Effective Unmanned Aerial Vehicle (UAV) Platform for Field Plant Breeding Application. *Remote Sens.* **2020**, *12*, 998. [[CrossRef](#)]
22. Themistocleous, K. The use of UAV platforms for remote sensing applications: Case studies in Cyprus. In Proceedings of the Second International Conference on Remote Sensing and Geoinformation of Environment, Pafos, Cyprus, 7–10 April 2014; Volume 92290S. [[CrossRef](#)]
23. Yang, G.; Liu, J.; Zhao, C.; Li, Z.; Huang, Y.; Yu, H.; Xu, B.; Yang, X.; Zhu, D.; Zhang, X.; et al. Unmanned Aerial Vehicle Remote Sensing for Field-Based Crop Phenotyping: Current Status and Perspectives. *Front. Plant Sci.* **2017**, *8*, 1111. [[CrossRef](#)]
24. Vergouw, B.; Nagel, H.; Bondt, G.; Custers, B. Drone Technology: Types, Payloads, Applications, Frequency Spectrum Issues and Future Developments. In *The Future of Drone Use*; Custers, B., Ed.; TMC Asser Press: The Hague, The Netherlands, 2016; pp. 3–20. [[CrossRef](#)]
25. Delavarpour, N.; Koparan, C.; Nowatzki, J.; Bajwa, S.; Sun, X. A Technical Study on UAV Characteristics for Precision Agriculture Applications and Associated Practical Challenges. *Remote Sens.* **2021**, *13*, 1204. [[CrossRef](#)]
26. Miphokasap, P.; Wannasiri, W. Estimations of Nitrogen Concentration in Sugarcane Using Hyperspectral Imagery. *Sustainability* **2018**, *10*, 1266. [[CrossRef](#)]
27. Wu, C.; Niu, Z.; Tang, Q.; Huang, W.; Rivard, B.; Feng, J. Remote estimation of gross primary production in wheat using chlorophyll-related vegetation indices. *Agric. For. Meteorol.* **2009**, *149*, 1015–1021. [[CrossRef](#)]
28. Lu, S.; Lu, X.; Zhao, W.; Liu, Y.; Wang, Z.; Omasa, K. Comparing vegetation indices for remote chlorophyll measurement of white poplar and Chinese elm leaves with different adaxial and abaxial surfaces. *J. Exp. Bot.* **2015**, *66*, 5625–5637. [[CrossRef](#)]
29. Fawcett, D.; Panigada, C.; Tagliabue, G.; Boschetti, M.; Celesti, M.; Evdokimov, A.; Biriukova, K.; Colombo, R.; Miglietta, F.; Rascher, U.; et al. Multi-Scale Evaluation of Drone-Based Multispectral Surface Reflectance and Vegetation Indices in Operational Conditions. *Remote Sens.* **2020**, *12*, 514. [[CrossRef](#)]
30. Osco, L.P.; Ramos, A.P.M.; Pereira, D.R.; Moriya, A.S.; Imai, N.N.; Matsubara, E.T.; Estrabis, N.; De Souza, M.; Junior, J.M.; Gonçalves, W.N.; et al. Predicting Canopy Nitrogen Content in Citrus-Trees Using Random Forest Algorithm Associated to Spectral Vegetation Indices from UAV-Imagery. *Remote Sens.* **2019**, *11*, 2925. [[CrossRef](#)]
31. Cui, B.; Zhao, Q.; Huang, W.; Song, X.; Ye, H.; Zhou, X. A New Integrated Vegetation Index for the Estimation of Winter Wheat Leaf Chlorophyll Content. *Remote Sens.* **2019**, *11*, 974. [[CrossRef](#)]

32. Gitelson, A.A.; Viña, A.; Ciganda, V.; Rundquist, D.C.; Arkebauer, T.J. Remote estimation of canopy chlorophyll content in crops. *Geophys. Res. Lett.* **2005**, *32*, L08403. [[CrossRef](#)]
33. Ballester, C.; Brinkhoff, J.; Quayle, W.C.; Hornbuckle, J. Monitoring the Effects of Water Stress in Cotton using the Green Red Vegetation Index and Red Edge Ratio. *Remote Sens.* **2019**, *11*, 873. [[CrossRef](#)]
34. Chen, P.; Haboudane, D.; Tremblay, N.; Wang, J.; Vigneault, P.; Li, B. New spectral indicator assessing the efficiency of crop nitrogen treatment in corn and wheat. *Remote Sens. Environ.* **2010**, *114*, 1987–1997. [[CrossRef](#)]
35. De Rosa, D.; Basso, B.; Fasiolo, M.; Friedl, J.; Fulkerson, B.; Grace, P.R.; Rowlings, D.W. Predicting pasture biomass using a statistical model and machine learning algorithm implemented with remotely sensed imagery. *Comput. Electron. Agric.* **2020**, *180*, 105880. [[CrossRef](#)]
36. Feng, L.; Zhang, Z.; Ma, Y.; Du, Q.; Williams, P.; Drewry, J.; Luck, B. Alfalfa Yield Prediction Using UAV-Based Hyperspectral Imagery and Ensemble Learning. *Remote Sens.* **2020**, *12*, 2028. [[CrossRef](#)]
37. Zhou, X.; Yang, L.; Wang, W.; Chen, B. UAV Data as an Alternative to Field Sampling to Monitor Vineyards Using Machine Learning Based on UAV/Sentinel-2 Data Fusion. *Remote Sens.* **2021**, *13*, 457. [[CrossRef](#)]
38. Lamichhane, S.; Kumar, L.; Wilson, B. Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: A review. *Geoderma* **2019**, *352*, 395–413. [[CrossRef](#)]
39. Fu, Y.; Yang, G.; Song, X.; Li, Z.; Xu, X.; Feng, H.; Zhao, C. Improved Estimation of Winter Wheat Aboveground Biomass Using Multiscale Textures Extracted from UAV-Based Digital Images and Hyperspectral Feature Analysis. *Remote Sens.* **2021**, *13*, 581. [[CrossRef](#)]
40. Han, L.; Yang, G.; Dai, H.; Xu, B.; Yang, H.; Feng, H.; Li, Z.; Yang, X. Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. *Plant Methods* **2019**, *15*, 10. [[CrossRef](#)]
41. Moran, J.A.; Mitchell, A.K.; Goodmanson, G.; Stockburger, K. Differentiation among effects of nitrogen fertilization treatments on conifer seedlings by foliar reflectance: A comparison of method. *Tree Physiol.* **2000**, *20*, 1113–1120. [[CrossRef](#)] [[PubMed](#)]
42. Xu, J.-X.; Ma, J.; Tang, Y.-N.; Wu, W.-X.; Shao, J.-H.; Wu, W.-B.; Wei, S.-Y.; Liu, Y.-F.; Wang, Y.-C.; Guo, H.-Q. Estimation of Sugarcane Yield Using a Machine Learning Approach Based on UAV-LiDAR Data. *Remote Sens.* **2020**, *12*, 2823. [[CrossRef](#)]
43. Canata, T.; Wei, M.; Maldaner, L.; Molin, J. Sugarcane Yield Mapping Using High-Resolution Imagery Data and Machine Learning Technique. *Remote Sens.* **2021**, *13*, 232. [[CrossRef](#)]
44. 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](#)]
45. QGIS.org. QGIS Geographic Information System. QGIS Association. 2021. Available online: <http://www.qgis.org> (accessed on 21 January 2022).
46. Imran, A.B.; Khan, K.; Ali, N.; Ahmad, N.; Ali, A.; Shah, K. Narrow band based and broadband derived vegetation indices using Sentinel-2 Imagery to estimate vegetation biomass. *Glob. J. Environ. Sci. Manag.* **2020**, *6*, 97–108. [[CrossRef](#)]
47. Marcial-Pablo, M.D.J.; Gonzalez-Sanchez, A.; Jimenez-Jimenez, S.I.; Ontiveros-Capurata, R.E.; Ojeda-Bustamante, W. Estimation of vegetation fraction using RGB and multispectral images from UAV. *Int. J. Remote Sens.* **2018**, *40*, 420–438. [[CrossRef](#)]
48. Avola, G.; Di Gennaro, S.F.; Cantini, C.; Riggi, E.; Muratore, F.; Tornambè, C.; Matese, A. Remotely Sensed Vegetation Indices to Discriminate Field-Grown Olive Cultivars. *Remote Sens.* **2019**, *11*, 1242. [[CrossRef](#)]
49. Boiarskii, B. Comparison of NDVI and NDRE Indices to Detect Differences in Vegetation and Chlorophyll Content. *J. Mech. Contin. Math. Sci.* **2019**, *4*, 20–29. [[CrossRef](#)]
50. Eitel, J.U.H.; Vierling, L.A.; Litvak, M.E.; Long, D.S.; Schulthess, U.; Ager, A.A.; Krofcheck, D.J.; Stoscheck, L. Broadband, red-edge information from satellites improves early stress detection in a New Mexico conifer woodland. *Remote Sens. Environ.* **2011**, *115*, 3640–3646. [[CrossRef](#)]
51. Zhang, J.; Wang, C.; Yang, C.; Xie, T.; Jiang, Z.; Hu, T.; Luo, Z.; Zhou, G.; Xie, J. Assessing the Effect of Real Spatial Resolution of In Situ UAV Multispectral Images on Seedling Rapeseed Growth Monitoring. *Remote Sens.* **2020**, *12*, 1207. [[CrossRef](#)]
52. Yu, R.; Luo, Y.; Zhou, Q.; Zhang, X.; Wu, D.; Ren, L. Early detection of pine wilt disease using deep learning algorithms and UAV-based multispectral imagery. *For. Ecol. Manag.* **2021**, *497*, 119493. [[CrossRef](#)]
53. Haboudane, D.; Miller, J.R.; Pattey, E.; Zarco-Tejada, P.J.; Strachan, I.B. Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sens. Environ.* **2004**, *90*, 337–352. [[CrossRef](#)]
54. Kumar, V.; Sharma, A.; Bhardwaj, R.; Thukral, A.K. Comparison of different reflectance indices for vegetation analysis using Landsat-TM data. *Remote Sens. Appl. Soc. Environ.* **2018**, *12*, 70–77. [[CrossRef](#)]
55. Gitelson, A.A.; Viña, A.; Arkebauer, T.J.; Rundquist, D.C.; Keydan, G.; Leavitt, B. Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophys. Res. Lett.* **2003**, *30*, 1248. [[CrossRef](#)]
56. Wu, W. The Generalized Difference Vegetation Index (GDVI) for Dryland Characterization. *Remote Sens.* **2014**, *6*, 1211–1233. [[CrossRef](#)]
57. Scher, C.L.; Karimi, N.; Glasenhardt, M.; Tuffin, A.; Cannon, C.H.; Scharenbroch, B.C.; Hipp, A.L. Application of remote sensing technology to estimate productivity and assess phylogenetic heritability. *Appl. Plant Sci.* **2020**, *8*, e11401. [[CrossRef](#)]
58. Jannoura, R.; Brinkmann, K.; Uteau, D.; Bruns, C.; Joergensen, R.G. Monitoring of crop biomass using true colour aerial photographs taken from a remote controlled hexacopter. *Biosyst. Eng.* **2015**, *129*, 341–351. [[CrossRef](#)]

59. Xue, J.; Su, B. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *J. Sens.* **2017**, *2017*, 1353691. [[CrossRef](#)]
60. Susantoro, T.M.; Wikantika, K.; Saepuloh, A.; Harsolumakso, A.H. Selection of vegetation indices for mapping the sugarcane condition around the oil and gas field of North West Java Basin, Indonesia. *IOP Conf. Ser. Earth Environ. Sci.* **2018**, *149*, 012001. [[CrossRef](#)]
61. Capolupo, A.; Monterisi, C.; Tarantino, E. Landsat Images Classification Algorithm (LICA) to Automatically Extract Land Cover Information in Google Earth Engine Environment. *Remote Sens.* **2020**, *12*, 1201. [[CrossRef](#)]
62. Kerkech, M.; Hafiane, A.; Canals, R. Deep learning approach with colorimetric spaces and vegetation indices for vine diseases detection in UAV images. *Comput. Electron. Agric.* **2018**, *155*, 237–243. [[CrossRef](#)]
63. Tesfaye, A.A.; Awoke, B.G. Evaluation of the saturation property of vegetation indices derived from sentinel-2 in mixed crop-forest ecosystem. *Spat. Inf. Res.* **2021**, *29*, 109–121. [[CrossRef](#)]
64. Melillos, G.; Hadjimitsis, D.G. Using simple ratio (SR) vegetation index to detect deep man-made infrastructures in Cyprus. In Proceedings of the Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XXV, Online, 27 April–8 May 2020; Volume 114180E. [[CrossRef](#)]
65. Salisu, A.; Abubakar, H.; Abubakar, H. One Way Anova: Concepts and Application in Agricultural System. In Proceedings of the CEUR Workshop Proceedings, Kaunas, Lithuania, 27 April 2018.
66. Blachnik, M. Comparison of Various Feature Selection Methods in Application to Prototype Best Rules. *Adv. Intell. Soft Comput.* **2009**, *57*, 257–264. [[CrossRef](#)]
67. Ramos, A.P.M.; Osco, L.P.; Furuya, D.E.G.; Gonçalves, W.N.; Santana, D.C.; Teodoro, L.P.R.; Junior, C.A.D.S.; Capristo-Silva, G.F.; Li, J.; Baio, F.H.R.; et al. A random forest ranking approach to predict yield in maize with uav-based vegetation spectral indices. *Comput. Electron. Agric.* **2020**, *178*, 105791. [[CrossRef](#)]
68. Catalina, T.; Iordache, V.; Caracaleanu, B. Multiple regression model for fast prediction of the heating energy demand. *Energy Build.* **2013**, *57*, 302–312. [[CrossRef](#)]
69. Perlich, C.; Provost, F.; Simonoff, J.S. Tree Induction vs. Logistic Regression: A Learning-Curve Analysis. *J. Mach. Learn. Res.* **2003**, *4*, 1–25.
70. Meek, C.; Thiesson, B.; Heckerman, D. The Learning-Curve Sampling Method Applied to Model-Based Clustering. *J. Mach. Learn. Res.* **2002**, *2*, 397–418.
71. Jozdani, S.E.; Johnson, B.A.; Chen, D. Comparing Deep Neural Networks, Ensemble Classifiers, and Support Vector Machine Algorithms for Object-Based Urban Land Use/Land Cover Classification. *Remote Sens.* **2019**, *11*, 1713. [[CrossRef](#)]
72. Wei, L.; Wang, Z.; Huang, C.; Zhang, Y.; Wang, Z.; Xia, H.; Cao, L. Transparency Estimation of Narrow Rivers by UAV-Borne Hyperspectral Remote Sensing Imagery. *IEEE Access* **2020**, *8*, 168137–168153. [[CrossRef](#)]
73. Yang, X.; Yang, R.; Ye, Y.; Yuan, Z.; Wang, D.; Hua, K. Winter wheat SPAD estimation from UAV hyperspectral data using cluster-regression methods. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *105*, 102618. [[CrossRef](#)]
74. Yoosefzadeh-Najafabadi, M.; Earl, H.J.; Tulpan, D.; Sulik, J.; Eskandari, M. Application of Machine Learning Algorithms in Plant Breeding: Predicting Yield from Hyperspectral Reflectance in Soybean. *Front. Plant Sci.* **2021**, *11*, 2169. [[CrossRef](#)]
75. Dong, T.; Shang, J.; Chen, J.M.; Liu, J.; Qian, B.; Ma, B.; Morrison, M.J.; Zhang, C.; Liu, Y.; Shi, Y.; et al. Assessment of Portable Chlorophyll Meters for Measuring Crop Leaf Chlorophyll Concentration. *Remote Sens.* **2019**, *11*, 2706. [[CrossRef](#)]
76. Liu, Y.; Liu, S.; Li, J.; Guo, X.; Wang, S.; Lu, J. Estimating biomass of winter oilseed rape using vegetation indices and texture metrics derived from UAV multispectral images. *Comput. Electron. Agric.* **2019**, *166*, 105026. [[CrossRef](#)]