



Article

Indoor Plant Soil-Plant Analysis Development (SPAD) Prediction Based on Multispectral Indices and Soil Electroconductivity: A Deep Learning Approach

Dorijan Radočaj * , Irena Rapčan and Mladen Jurišić

Faculty of Agrobiotechnical Sciences Osijek, Josip Juraj Strossmayer University of Osijek, Vladimira Preloga 1, 31000 Osijek, Croatia; irapcan@fazos.hr (I.R.); mjurisic@fazos.hr (M.J.)

* Correspondence: dradocaj@fazos.hr; Tel.: +385-31-554-965

Abstract: Leaf Soil-Plant Analysis Development (SPAD) prediction is a crucial measure of plant health and is essential for optimizing indoor plant management. The deep learning methods offer advanced tools for precise evaluations but their adaptation to the heterogeneous indoor plant ecosystem presents distinct challenges. This study assesses how accurately deep neural network (DNN) predicts SPAD values in leaves on indoor plants when compared to well-established machine learning techniques, including Random Forest (RF) and Extreme Gradient Boosting (XGB). The covariates for prediction were based on low-cost multispectral and soil electro-conductivity (EC) sensors, enabling a non-destructive sensing approach. The study also strongly emphasized multicollinearity analysis quantified by the Variance Inflation Factor (VIF) and two independent indices, as well as its effect on prediction accuracy using deep and machine learning methods. DNN resulted in higher accuracy to RF and XGB, also performing better using filtered data after multicollinearity analysis based on the coefficient of determination (R^2), root mean square error (RMSE) and mean absolute error (MAE) ($R^2 = 0.589$, RMSE = 11.68, MAE = 9.52) in comparison to using all input covariates ($R^2 = 0.476$, RMSE = 12.90, MAE = 10.94). Overall, DNN was proven as a more accurate prediction method than the conventional machine learning approach for the prediction of leaf SPAD values in indoor plants, despite using heterogenous plant types and input covariates.

Keywords: deep neural networks; regression; chlorophyll concentration; Plant-O-Meter; multicollinearity analysis



Citation: Radočaj, D.; Rapčan, I.; Jurišić, M. Indoor Plant Soil-Plant Analysis Development (SPAD) Prediction Based on Multispectral Indices and Soil Electroconductivity: A Deep Learning Approach. *Horticulturae* **2023**, *9*, 1290. <https://doi.org/10.3390/horticulturae9121290>

Academic Editors: Alessia Cogato, Marco Sozzi, Eve Laroche-Pinel and Nebojša Nikolić

Received: 4 November 2023
Revised: 21 November 2023
Accepted: 29 November 2023
Published: 30 November 2023



Copyright: © 2023 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

In an increasingly fast-paced and technologically driven world, where the majority of working hours are spent indoors, the importance of creating a work environment that is both conducive and sustainable cannot be overstated [1]. The modern workplace is more than just a physical place to perform tasks; it is now a dynamic ecosystem with the ability to dramatically affect the physical and psychological well-being of its occupants [2]. The introduction of greenery into interior spaces is not only motivated by aesthetic preferences but is also recognized as a strategic requirement in the quest to create workplaces that promote employee health and productivity [3]. More than just decorative elements, indoor plants have been shown to improve air quality, reduce stress, increase productivity, and contribute to a more desirable and productive workplace [4]. This increased awareness of the many benefits of indoor plants has created new opportunities for research and innovation in the fields of agriculture and horticulture. Precise assessment and prediction of plant health and vitality have emerged as critical variables in ensuring optimal crop development and effective resource management in the context of modern horticulture [5,6]. Measurement of chlorophyll concentration is a key indicator used to assess overall plant health due to its direct relationship with photosynthetic activity [7]. In addition, phytoremediation, the use of plants to remove toxins from indoor environments, is becoming increasingly

important [8]. In such cases, accurate prediction of chlorophyll concentration helps select indoor plant species with the greatest potential to effectively clean indoor air [9]. The ability to differentiate between plants based on chlorophyll concentration can greatly improve the efficiency of phytoremediation efforts, making indoor environments healthier and more sustainable.

Chlorophyll content measurements are frequently represented by Soil-Plant Analysis Development (SPAD) values, which have become essential tools for assessing plant health [10]. Variables that affect SPAD meters include the surrounding environment, leaf age, species-specific characteristics, and nutrient availability [11]. Standardized techniques to improve reliability have been proposed through calibration and validation studies that have examined species-specific calibration curves and the effect of changing conditions on SPAD meter performance. However, many of these studies proposed calibration equations that are empirical and based on a single crop type, meaning that these are not viable for often highly heterogeneous indoor plant types [12–14]. Studies contrasting SPAD values with conventional techniques highlight the relationship between these values and indices of plant health and a very high correlation with leaf chlorophyll content [15]. A thorough understanding of SPAD estimation is becoming increasingly important as indoor plant management requires optimizing development and maintaining plants efficiently. As research into the process of predicting SPAD values has progressed, new advances in deep learning approaches are providing tools for ever more accurate assessments [16,17]. However, previous research has mostly focused on outdoor plants, so applying these methods to indoor plants presents a unique set of obstacles. A wide variety of plant species and variable abiotic factors characterize the indoor environment, resulting in increased biodiversity [18]. As a result, using deep learning techniques to reliably predict SPAD values in indoor plants requires novel approaches and specialized models adapted to this distinct and dynamic environment. Previous studies have highlighted the critical importance of integrating complementary and independent sensors in the context of automated plant health monitoring via the Internet of Things (IoT) [19]. The integration of such sensor technology offers a novel paradigm for accurate and real-time plant health monitoring in the era of smart and sustainable indoor environments [20]. Multispectral sensors collect spectral data at multiple wavelengths, enabling non-destructive inspection of plant leaves and, as a result, more accurate estimates of SPAD values [21]. Soil electroconductivity (EC) measurements, on the other hand, provide an indirect but important indication of soil nutrient levels and general environmental conditions [22]. These sensor modalities, when combined, provide an in-depth view of the plant-soil system, providing information on the health and development potential of indoor plants. Several authors used soil EC directly as an important physico-chemical-biological component of plant health [23], as well as an indirect representation of soil moisture availability for plant development in a broad scope of studies [24]. In the context of the IoT, these sensors enable automated and continuous monitoring, providing real-time input that can be used for timely intervention and efficient resource management [25]. Deep and machine learning approaches use sensor data to improve prediction accuracy, fostering the development of intelligent systems capable of maintaining thriving indoor plant ecosystems, with broader implications for occupant well-being and indoor sustainability [26].

The advanced search of studies indexed in the Web of Science Core Collection with the broad topics of “SPAD” and “deep learning”, as well as “SPAD” and “indoor plants” returned 27 and 10 studies, respectively. While present research which used deep learning exclusively used hyperspectral and multispectral images for outdoor arable crops [27–30], those focusing on indoor plants used SPAD values as auxiliary indicators of the effectiveness of LED light in an indoor environment [31,32]. Therefore, there is currently no research that provides indoor plant health monitoring quantified by SPAD values which is suitable for the implementation in the IoT. To address this research gap in predicting leaf SPAD values of indoor plants based on state-of-the-art deep learning, the main objective of this study was to evaluate its efficiency compared to well-known machine learning methods. Furthermore,

the data collection for covariates was based on non-destructive multispectral and soil EC sensors, allowing the proposed procedure to be implemented as part of automated plant monitoring in the IoT.

2. Materials and Methods

The data collection in this study comprised two primary components: (1) leaf SPAD measurement of indoor plants as training and test data for prediction; and (2) leaf multi-spectral and soil EC sensing for modeling covariates for the prediction. The deep neural network (DNN) was proposed for leaf SPAD prediction and was evaluated alongside two well-known machine learning methods.

2.1. Indoor Plants Analyzed in the Study

A total of 52 individual indoor plants of ten species were analyzed in the study, as shown in Figure 1. Data collection was performed on 13 October 2023 in the building of the Faculty of Agrobiotechnical Sciences Osijek, which covers an area of 18,600 m². The studied indoor plants were located on the first three floors of the building, which were distributed on the four main sides of the building, as well as in two central corridors connecting the north and south sides of the building. All indoor plants were maintained in standardized containers as shown in Figure 1, and were also watered, fertilized, and managed in a standardized manner throughout the building.



Figure 1. The display of ten indoor plant species analyzed in the study, maintained in standardized containers at the Faculty of Agrobiotechnical Sciences Osijek.

2.2. Sensors and Sensing Approach Used for Plant and Soil Measurement

Three non-destructive sensing approaches were used to model leaf SPAD values, as well as plant and soil covariates for deep and machine learning prediction of leaf SPAD (Figure 2). Furthermore, all sensors used are low-cost solutions, which allows their widespread implementation, both as standalone systems [33] or as part of the IoT [34,35].

The SPAD meter of indoor plant leaves was measured using the Konica Minolta SPAD-502 Plus handheld chlorophyll sensor (Tokyo, Japan). SPAD values were used to represent relative leaf chlorophyll content based on absorbance measurement, achieving a very high correlation with leaf nitrogen concentration [15]. SPAD measurements per plant were taken as an average of six measurements at evenly spaced points, taking into account present variations in leaf color and condition in proportion to overall plant condition.

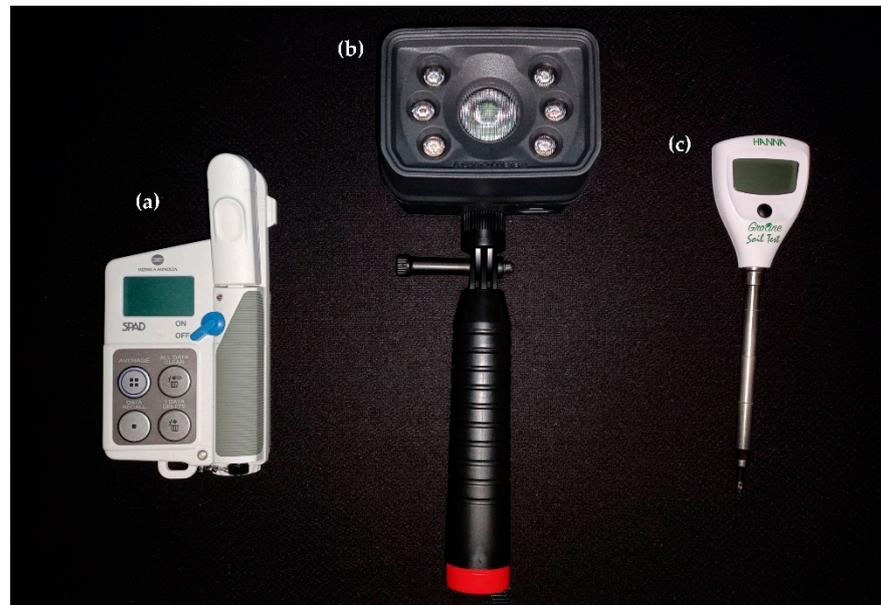


Figure 2. Low-cost handheld sensors used in the study: (a) Konica Minolta SPAD-502 Plus chlorophyll sensor, (b) multispectral Plant-O-Meter sensor (Novi Sad, Serbia), (c) Hanna instruments HI 98331 soil EC sensor.

The covariates for predicting SPAD values for indoor plants were collected using two different approaches based on plant and soil sensing: (1) multispectral sensing of plant leaves and (2) soil EC measurement (Figure 3). Multispectral sensing of the indoor plant canopy was performed using the Plant-O-Meter, a handheld device that operates in five wavelengths (blue, green, red, red-edge, and near-infrared bands) [36]. All measurements were collected and exported using the Plant-O-Meter Android application connected to the handheld device via Bluetooth. Because Plant-O-Meter operates as an active sensor, accurate canopy sensing was maintained regardless of lighting conditions. A total of 21 vegetation indices were calculated based on the Plant-O-Meter measurements, as listed in Table 1. These indices quantified a wide range of vegetation properties, providing a basis for implementing deep and machine-learning methods in a variety of agricultural studies [37]. Because the indoor plant species evaluated in this study differed by canopy system, measurements were made by aiming the sensor with the criterion of covering a plane to include maximum leaf coverage during sampling. Soil EC was sampled using the Hanna Instruments HI 98331 Handheld Soil EC Sensor (Nusfalau, Romania) at 5 cm, 10 cm, and 15 cm soil depths within a 10 cm radius of plant stems, allowing for non-destructive measurement toward root systems. This resulted in a total of 52 input samples and 25 covariates evaluated in the study.

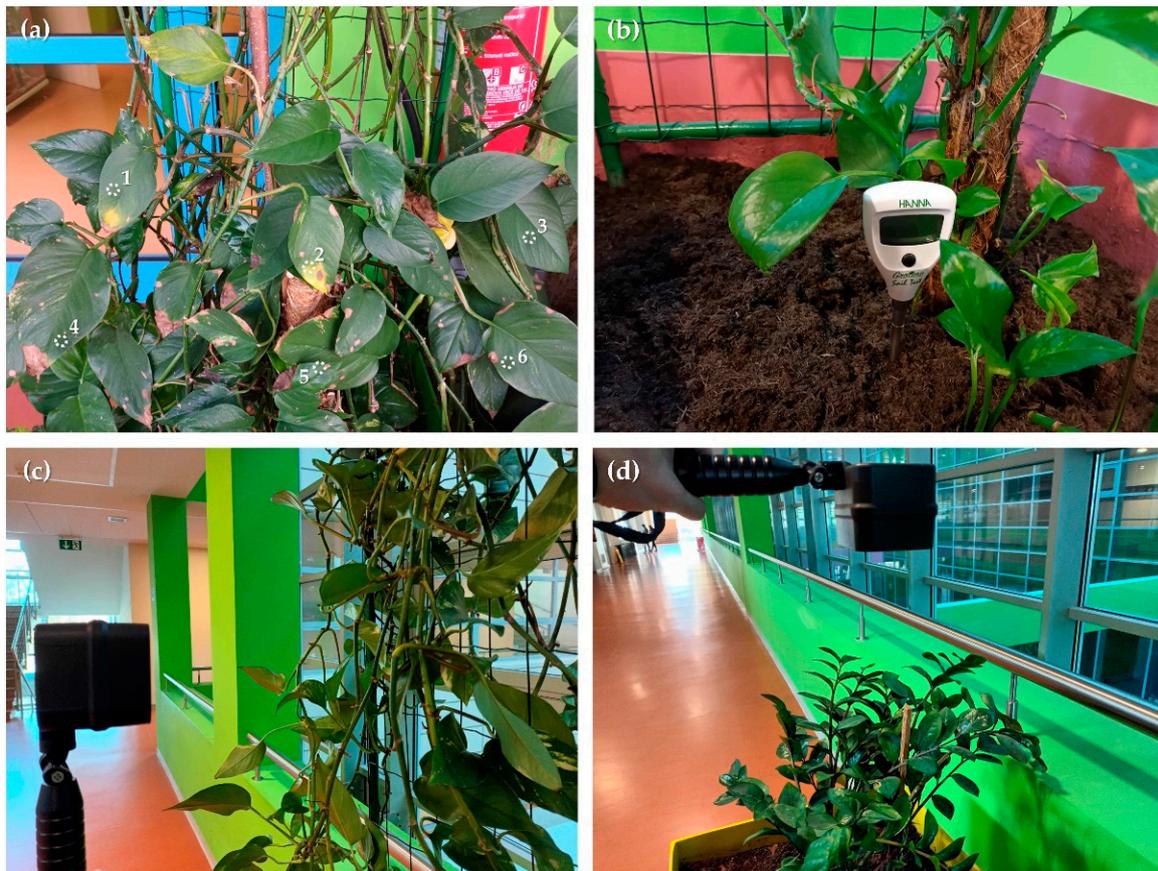


Figure 3. The display of sensing process for: (a) six evenly distributed and representative points per plant using chlorophyll sensor, (b) soil EC measurement, (c,d) multispectral sensing using Plant-O-Meter, based on plant canopy system.

Table 1. Vegetation indices collected using Plant-O-Meter as covariates for the prediction of leaf SPAD values.

Vegetation Index	Formula	Reference
Red normalized difference vegetation index	$NDVIr = \frac{NIR - R}{NIR + R}$	[38]
Green normalized difference vegetation index	$NDVIg = \frac{NIR - G}{NIR + G}$	[39]
Blue normalized difference vegetation index	$NDVIb = \frac{NIR - B}{NIR + B}$	[40]
Normalized difference red-edge vegetation index	$NDRE = \frac{NIR - RE}{NIR + RE}$	[41]
Simple ratio	$SR = \frac{NIR}{R}$	[42]
Dark green color index	$DGCI = \left[\frac{(hue-60)}{60} + (1-saturation) + (1-brightness) \right]$	[43]
Structure insensitive pigment index	$SIPI = \frac{NIR - B}{NIR + R}$	[44]
Red-green ratio	$RGR = \frac{R}{G}$	[45]
Modified simple ratio	$MSR = \frac{R}{\sqrt{\frac{NIR}{R} + 1}}$	[46]
Renormalized difference vegetation index	$RDVI = \frac{NIR - R}{\sqrt{NIR - R}}$	[47]
Infrared percentage Vegetation index	$IPVI = \frac{NIR}{NIR + R}$	[48]
Enhanced vegetation index	$EVI = \frac{2.5(NIR - R)}{NIR + 6R - 7.5B + 1}$	[49]
Green soil adjusted vegetation index	$GSAVI = 1.5 \frac{NIR - G}{NIR + G + 0.5}$	[50]
Green optimized soil adjusted vegetation index	$GOSAVI = \frac{NIR - G}{NIR + G + 0.16}$	[50]
Transformed difference vegetation index	$TDVI = 1.5 \frac{NIR - R}{\sqrt{NIR^2 - R^2 + 0.5}}$	[51]

Table 1. Cont.

Vegetation Index	Formula	Reference
Wide dynamic range vegetation index	$WDRVI = \frac{0.28 \text{ NIR} - \text{R}}{0.28 \text{ NIR} + \text{R}}$	[52]
Green-red normalized difference vegetation index	$GRNDVI = \frac{\text{NIR} - (\text{G} + \text{R})}{\text{NIR} + \text{G} + \text{R}}$	[40]
Green-blue normalized difference vegetation index	$GBNDVI = \frac{\text{NIR} - (\text{G} + \text{B})}{\text{NIR} + \text{G} + \text{B}}$	[40]
Red-blue normalized difference vegetation index	$RBNDVI = \frac{\text{NIR} - (\text{R} + \text{B})}{\text{NIR} + \text{R} + \text{B}}$	[40]
Visible normalized difference vegetation index	$PNDVI = \frac{\text{NIR} - (\text{G} + \text{R} + \text{B})}{\text{NIR} + \text{G} + \text{R} + \text{B}}$	[40]
Inverted simple ratio	$ISR = \frac{\text{R}}{\text{NIR}}$	[53]

B: reflectance in blue band (455 nm), G: reflectance in green band (528 nm), R: reflectance in red band (657 nm), RE: reflectance in red-edge band (740 nm), NIR: reflectance in near-infrared band (810 nm).

Variance Inflation Factor (VIF) analysis was used to evaluate all multispectral and soil EC covariates for multicollinearity. The VIF was used as the primary metric to measure the degree of multicollinearity among predictor variables in a regression model. The VIF values generated for each covariate indicated the extent to which they were related, with values greater than 10 strongly indicating multicollinearity [54]. In addition, two multicollinearity indices (IND1 and IND2) proposed by Ullah et al. [55] were used independently to assess multicollinearity. IND1 values greater than 0 and IND2 values less than 1 indicated less multicollinearity.

2.3. Deep and Machine Learning Prediction and Accuracy Assessment

The DNN was built sequentially using the Keras library in R v4.0.3, with each layer added sequentially to facilitate the flow of data from input to output, as shown in Figure 4. The data were normalized before entering the network using the -1L axis normalization, which indicates normalization over the last axis. The normalizer was then fitted to the training features by applying the fit function to the training data, which was represented as a matrix. The network consists of several dense (fully connected) layers, each with its own set of features. The first layer consisted of 32 units and used the Rectified Linear Unit (ReLU) activation function, which is a popular choice for hidden layers [56,57]. In addition, to avoid overfitting, these layers were subjected to L1 regularization with a regularization strength of 0.001. To incorporate regularization by randomly deactivating a percentage of input units during training, a dropout layer with a dropout rate of 0.1 was used. This was followed by another 32-unit dense layer with ReLU activation, followed by a pair of 16-unit layers, the first with L1 regularization and the second without explicit regularization. Finally, the output layer was a single-unit dense layer, since the network was designed for a regression problem. The overall design incorporates many features aimed at maximizing predictive performance and controlling model complexity for the problem domain addressed in this study.

Two machine learning methods, Random Forest (RF) and Extreme Gradient Boosting (XGB), were evaluated alongside DNN. RF and XGB achieved superior prediction accuracy in regression problems compared to current machine learning algorithms in similar studies on various aspects of horticulture [58,59] and agriculture in general [60–62]. As an ensemble learning technique, RF builds a forest of decision trees, each trained separately on randomly selected samples of the data and features [63]. Averaging the predictions from these trees reduces overfitting, improves model robustness, and captures complex interactions between variables and the target variable. In contrast, the XGB gradient boosting technique uses an iterative process based on decision trees [64]. Starting with a simple model, it builds decision trees in stages, each of which aims to minimize the errors produced by the previous iterations. Through this repeated learning process, XGB can adapt to complicated data relationships and continuously improve predictions. This process is guided by the calculation of residuals. Together, these two techniques successfully take advantage of both the ensemble approach of RF and the iterative improvement of XGB, enabling the model to estimate leaf SPAD values with high accuracy. A tuning hyperparameter for RF

was the quantity of variables randomly sampled at each split (*mtry*), while for XGB the contribution of each tree to the overall ensemble was affected by the learning rate (*eta*), the number of boosting rounds (*nrounds*), while *alpha* controlled the L1 regularization and *lambda* controlled the L2 regularization on the leaf weights. The hyperparameter turning was performed using a built-in automated approach in *caret* library.

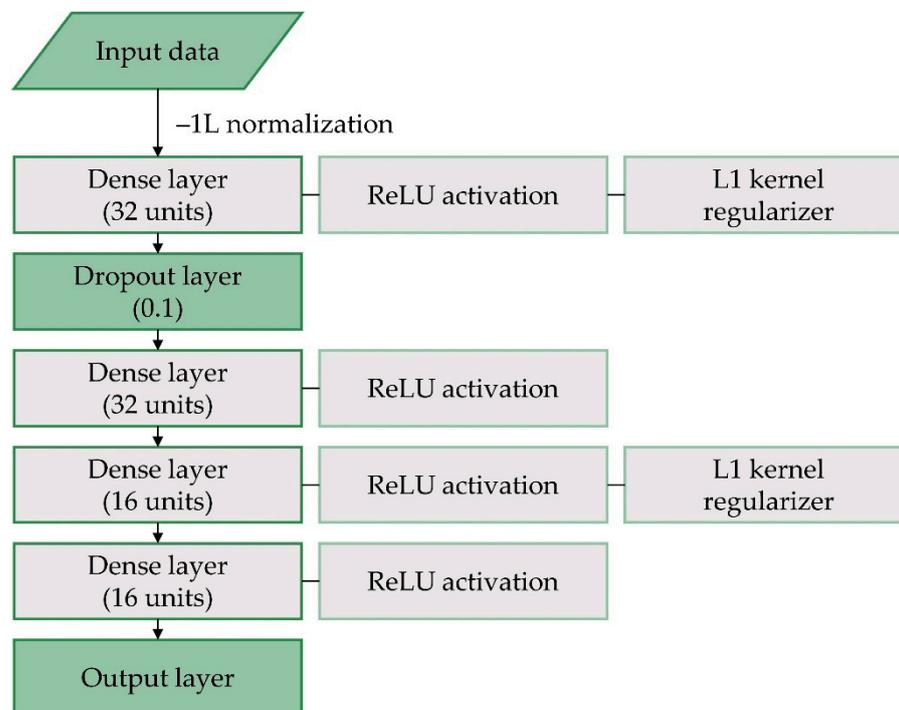


Figure 4. The architecture of the proposed DNN for prediction of leaf SPAD values of indoor plants.

DNN, RF, and XGB were evaluated in two approaches based on the input data, considering all multispectral and soil covariates (all input data) and only covariates for which multicollinearity was not detected (filtered input data). Three statistical measures were used to assess the accuracy and reliability of the prediction models using 10-fold cross-validation. The primary metric used was the coefficient of determination (R^2), which quantifies the proportion of variation in predicted SPAD explained by the model. A higher R^2 value indicates a better fit of the model to the data, indicating its ability to accurately capture the underlying relationships. In addition, the root mean square error (RMSE) and mean absolute error (MAE) were used to assess the prediction performance of the model. The average size of prediction errors is quantified by the RMSE, which provides a measure of how well the model's predictions match the observed data. The average size of the absolute errors, expressed as MAE, provides a more robust insight into the accuracy of the model. The use of R^2 , RMSE, and MAE in this accuracy assessment provided a comprehensive evaluation of leaf SPAD prediction and allowed a comprehensive analysis of the accuracy of DNN, RF, and XGB.

3. Results and Discussion

The Pearson's correlation coefficients shown in the correlation plot (Figure 5) indicate two main results: (1) the indoor plant type and soil EC covariates generally produced low correlations compared to the Plant-O-Meter vegetation indices, and (2) the individual vegetation indices tended to produce high and very high absolute Pearson's correlation coefficients when evaluated against each other. The exception to the latter was to some extent present for DGCI, RGR, RDVI, NDRE and EVI. DGCI is the most distinct vegetation index among those evaluated because it uses band values transformed to the hue-saturation-brightness model [65], resulting in Pearson's correlation coefficients with other vegetation

indices ranging from -0.73 to 0.48 . Although the RGR calculation is based on a simple red-green ratio, it produced weak to moderate correlations with other vegetation indices. This was probably due to the specific selection of Plant-O-Meter vegetation indices, which focused dominantly on the use of red and near-infrared bands [36], with RGR being the only evaluated index using only red and green bands. However, there is no basis for the relatively low positive correlation of RDVI with almost all other vegetation indices related to band selection. Similar to RGR, NDRE resulted in low to moderate correlations with other vegetation indices, ranging from -0.49 to 0.63 , because it was the only vegetation index evaluated that used the red edge band. Aside from its specificity in band selection, NDRE likely provided different results from the majority of indices based on red and near-infrared bands (especially NDVIr) due to its resistance to saturation effect in cases of high biomass [66]. Based on previous studies, EVI was expected to provide different results from vegetation indices using only red and near-infrared bands and likely provided a more robust assessment of plant health due to the inclusion of blue bands [67].

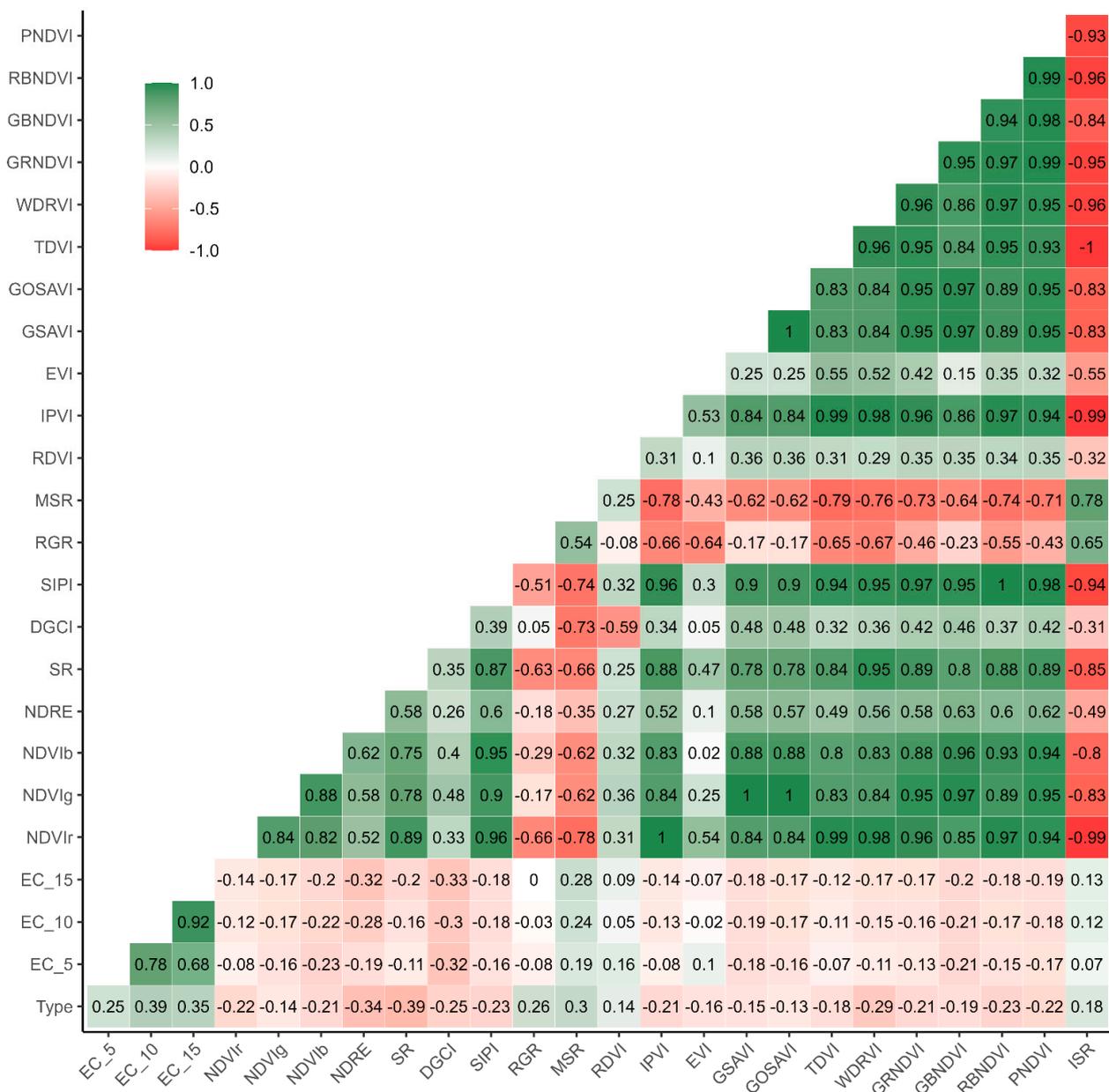


Figure 5. The correlation plot of all input covariates evaluated for the prediction of leaf SPAD of indoor plants, with values representing Pearson’s correlation coefficient.

The multicollinearity indices used in the study, VIF, IND1, and IND2, partially agreed with the results of the correlation analysis regarding the covariate correlation (Table 2). With the primary criterion of VIF values less than 10 [54], only four out of a total of 25 covariates indicated an absence of multicollinearity, including plant type, soil EC at 5 cm and 15 cm soil depth, and NDRE. While the results of plant type and soil EC measurements strongly indicated their independence relative to all input covariates, NDRE was the only vegetation index that resulted in the absence of multicollinearity, probably due to the previously mentioned ability of the red-edge band to provide resistance to saturation in cases of higher biomass [67]. The multicollinearity observations of VIF were confirmed by the IND1 and IND2 indices, providing an independent check of the multicollinearity analysis [55]. The final data filtered after the multicollinearity analysis consisted of four covariates, including plant type, soil EC at 5 cm and 15 cm soil depth, along with NDRE, as a single vegetation index.

Table 2. The results of multicollinearity analysis for all 25 input covariates evaluated in the study.

Variable	VIF	IND1	IND2
Type	3.2177	0.2763	0.7223
EC_5	5.3337	0.1667	0.8515
EC_10	14.9256	0.0596	0.9778
EC_15	9.2927	0.0957	0.9352
NDVlr	2492.506	0.0004	1.0476
NDVIg	6035.032	0.0001	1.0478
NDVIb	569.3946	0.0016	1.0462
NDRE	3.3274	0.2671	0.7330
SR	95.0013	0.0094	1.0370
DGCI	39.0306	0.0228	1.0211
SIPI	1699.364	0.0005	1.0474
RGR	235.9476	0.0038	1.0436
MSR	48.4881	0.0183	1.0264
RDVI	18.3978	0.0483	0.9910
IPVI	792.5418	0.0011	1.0467
EVI	20.537	0.0433	0.9970
GSAVI	3612.891	0.0002	1.0477
GOSAVI	6294.054	0.0001	1.0478
TDVI	2112.216	0.0004	1.0475
WDRVI	1432.165	0.0006	1.0473
GRNDVI	3869.734	0.0002	1.0477
GBNDVI	2864.549	0.0003	1.0476
RBNDVI	3019.629	0.0003	1.0477
PNDVI	4943.411	0.0002	1.0478
ISR	1661.771	0.0005	1.0474

Input variables which did not indicate multicollinearity are bolded.

Across all evaluation measures, the DNN model outperformed RF and XGB in terms of prediction accuracy with the primary criterion of higher R^2 (Table 3), while also achieving highly consistent prediction accuracy regardless of training and test data folds during cross-validation (Figure 6). The optimal hyperparameters considering input data were the same for RF (mtry = 2), while XGB produced higher regularization strength for all input data (eta = 0.3, nrounds = 50, lambda = 0.1, alpha = 0.0001), in comparison to filtered input data (eta = 0.3, nrounds = 100, lambda = 0, alpha = 0). Furthermore, it is noteworthy that the results using filtered input data after multicollinearity analysis improve the prediction accuracy for DNN, which was so far usually evaluated only for machine learning methods [68]. The two evaluated machine learning methods were less sensitive to input covariate selection, with XGB producing slightly higher prediction accuracy using filtered input data. These results underscore the need for multicollinearity analysis in similar studies that consider numerous input covariates with DNN, supporting and expanding on the observations of McCaw et al. [69], despite mixed results regarding

relative accuracy quantified by R^2 and absolute prediction accuracy represented by RMSE and MAE. Meanwhile, machine learning prediction results confirm the observations of previous studies in which was proven that RF can effectively handle multicollinearity, providing robust predictions despite a high correlation between input features [70], while XGB may be more prone to the effects of multicollinearity, given its tendency to construct deeper and more intricate trees [71]. The results of 10-fold cross-validation also strongly indicated its superiority over split-sample accuracy assessment, which greatly varies among folds for RF and XGB, and could otherwise imply inconsistent and unreliable prediction performance in heterogeneous input datasets [72]. Unlike more frequent studies that use deep learning based on images, Graditi et al. [73] noted that regression problems do not require a large amount of input data for accurate prediction, which was confirmed in studies based on both deep [74] and conventional machine learning [75]. To ensure resistance to overfitting during the predictions based on smaller datasets, Hosseini et al. [76] strongly recommended implementing k-fold cross-validation, as was performed in this study, instead of a simpler and more frequent split-sample approach. However, a study by Gilbertson and van Niekerk proved that, while smaller datasets can reliably produce moderately high prediction accuracy, the addition of a larger amount of training samples would likely lead to higher prediction accuracy [77].

Table 3. Prediction accuracy of evaluated deep and machine learning methods in leaf SPAD prediction of indoor plants.

Method	Value	All Input Data			Filtered Input Data		
		RF	XGB	DNN	RF	XGB	DNN
R^2	Mean	0.504	0.430	0.476	0.461	0.522	0.589
	CV	0.725	0.804	0.547	0.629	0.511	0.588
RMSE	Mean	10.65	12.82	12.90	11.30	13.15	11.68
	CV	0.458	0.302	0.238	0.516	0.510	0.521
MAE	Mean	8.35	10.36	10.94	9.07	10.39	9.52
	CV	0.471	0.263	0.216	0.492	0.465	0.541

The most accurate prediction metrics are bolded.

Although SPAD values are widely used, there are several limitations when it comes to representing the chlorophyll content of indoor plant leaves. One notable limitation is the potential effect of environmental variables on SPAD measurements since chlorophyll fluorescence can be affected by changes in temperature, humidity, and light intensity, which can introduce variability into SPAD readings [78]. In addition, there are difficulties in calibrating SPAD meters because the best calibration curves vary with species and highly heterogeneous indoor growing environments. In addition, because SPAD values cannot distinguish between chlorophyll a and chlorophyll b, they can only be used to estimate the amount of chlorophyll present [15]. The possible effect of nutrient supplementation on SPAD readings should be considered, as indoor plant habitats are often subject to regulated conditions and excessive or deficient nutrient levels may result in a false representation of chlorophyll status. Due to the considerable heterogeneity in input indoor plant samples having ten species represented by 52 samples, the proposed approach can be expected to be robust with similar indoor plant datasets. Moreover, the increased sample count while retaining a similar amount of plant species will likely result in an increased prediction accuracy in all evaluated instances [77].

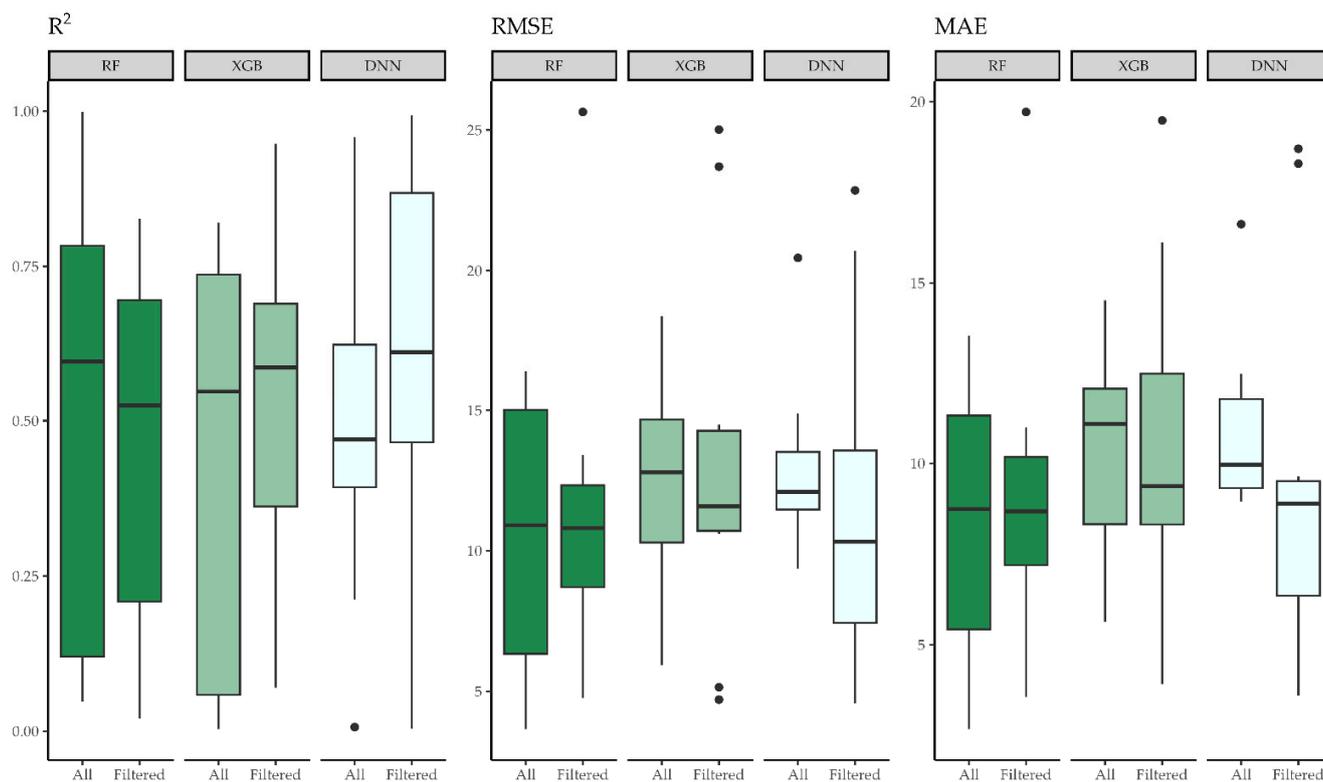


Figure 6. A display of variability of accuracy assessment metrics across 10 folds during cross-validation.

Vegetation indices are widely recognized as sensitive indicators of plant health due to their ability to capture various aspects of vegetation dynamics [37,79]. Thus, it is important to account for the effect of biomass when utilizing vegetation indices for plant health assessments [80]. However, the influence of biomass on these indices was not explicitly considered in the research. The lack of inclusion of biomass-related factors in the research affects the spectral reflectance properties [81], thus limiting the predictive accuracy of the model. To overcome this limitation, further calibration and refinement of the proposed approach are possible. Integration of this approach with deep learning-based plant identification algorithms also presents an avenue for future enhancement [82]. By utilizing the potential of vegetation indices and advanced plant species recognition technology based on deep learning, a more comprehensive and consistent model for predicting plant health can be developed. This integration may provide a more thorough understanding of vegetation status, including species-specific intricacies and overall biomass-related dynamics, consequently enhancing the robustness and dependability of plant health predictions.

4. Conclusions

The process of plant health assessment has changed significantly since deep learning was introduced to the field. However, there are particular difficulties in applying these methods to complex and heterogeneous indoor plant ecosystems. This research addressed these issues and highlighted the effectiveness of DNN in predicting leaf SPAD of indoor plants based on a non-destructive approach.

Indoor plant types and soil EC showed lower correlations compared to the vegetation indices obtained from the Plant-O-Meter. The evaluation of individual vegetation indices against each other resulted in high and very high absolute Pearson's correlation coefficients, except for DGCI, RGR, RDVI, NDRE, and EVI. Partial concurrence between multicollinearity analysis using VIF, IND1, and IND2 with the correlation results was observed. Only four out of the 25 covariates, which were plant type, soil EC at 5 cm and 15 cm soil depth, and NDRE, showed an absence of multicollinearity as indicated by VIF values. The only

vegetation index without detected multicollinearity was NDRE, likely due to the resistance of its red-edge band to saturation effects in high-biomass scenarios. The fact that VIF, IND1, and IND2 all show consistent results provides solid evidence of the accuracy of the analysis regarding multicollinearity.

The DNN model outperformed both RF and XGB in terms of predictive accuracy across several evaluation measures. Additionally, it showed superior consistency in its prediction accuracy across various training and testing data folds during cross-validation. Notably, this robustness suggests a dependable predictive performance in diverse and heterogeneous input datasets. Filtering the input data based on multicollinearity analysis improved the prediction accuracy of the DNN model, which highlights the importance of accounting for multicollinearity in similar studies. On the other hand, RF demonstrated the ability to handle multicollinearity effectively, providing robust predictions despite high correlations between input features, while XGB produced moderately high accuracy but was more susceptible to multicollinearity because it constructed deeper and more intricate trees.

Furthermore, by incorporating these technologies into the IoT framework, it is possible to automatically monitor plant health in real time, which promises to create healthier indoor environments. The results of this study underscore the importance of considering multicollinearity when using DNN by selecting variables and emphasize the need for accuracy in data collection.

Author Contributions: Conceptualization, D.R.; methodology, D.R.; software, D.R.; validation, D.R., I.R. and M.J.; formal analysis, D.R., I.R. and M.J.; investigation, D.R.; resources, D.R. and M.J.; data curation, D.R.; writing—original draft preparation, D.R.; writing—review and editing, D.R., I.R. and M.J.; visualization, D.R.; supervision, M.J.; project administration, M.J.; funding acquisition, D.R., I.R. and M.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the inclusion of information that could compromise the privacy of research participants.

Acknowledgments: The authors are grateful to Časlav Livada of the Faculty of Electrical Engineering, Computer Science and Information Technology Osijek for providing valuable suggestions on deep learning prediction method.

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

References

1. Kroemer, A.D.; Kroemer, K.H.E. *Office Ergonomics: Ease and Efficiency at Work*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2016; ISBN 978-1-4987-4797-4.
2. Altomonte, S.; Allen, J.; Bluysen, P.M.; Brager, G.; Hescong, L.; Loder, A.; Schiavon, S.; Veitch, J.A.; Wang, L.; Wargocki, P. Ten Questions Concerning Well-Being in the Built Environment. *Build. Environ.* **2020**, *180*, 106949. [[CrossRef](#)]
3. Elnaklah, R.; Fosas, D.; Natarajan, S. Indoor Environment Quality and Work Performance in “Green” Office Buildings in the Middle East. *Build. Simul.* **2020**, *13*, 1043–1062. [[CrossRef](#)]
4. Al Horr, Y.; Arif, M.; Kaushik, A.; Mazroei, A.; Katafygiotou, M.; Elsarrag, E. Occupant Productivity and Office Indoor Environment Quality: A Review of the Literature. *Build. Environ.* **2016**, *105*, 369–389. [[CrossRef](#)]
5. El-Ramady, H.R.; Alshaal, T.A.; Amer, M.; Domokos-Szabolcsy, É.; Elhawati, N.; Prokisch, J.; Fári, M. Soil Quality and Plant Nutrition. In *Sustainable Agriculture Reviews 14: Agroecology and Global Change*; Ozier-Lafontaine, H., Lesueur-Jannoyer, M., Eds.; Sustainable Agriculture Reviews; Springer International Publishing: Cham, Switzerland, 2014; pp. 345–447. ISBN 978-3-319-06016-3.
6. Lastochkina, O.; Aliniaiefard, S.; SeifiKalhor, M.; Bosacchi, M.; Maslennikova, D.; Lubyanova, A. Novel Approaches for Sustainable Horticultural Crop Production: Advances and Prospects. *Horticulturae* **2022**, *8*, 910. [[CrossRef](#)]
7. Bemejo-Padilla, E.; Kinsou, H.; Filali, R.; Perez-Bibbins, B.; Taidi, B. Rapid Indicators for Monitoring the Health of *Chlamydomonas Nivalis* Biomass during Preservation. *J. Appl. Phycol.* **2021**, *33*, 2723–2732. [[CrossRef](#)]
8. Jeevanantham, S.; Saravanan, A.; Hemavathy, R.V.; Kumar, P.S.; Yaashikaa, P.R.; Yuvaraj, D. Removal of Toxic Pollutants from Water Environment by Phytoremediation: A Survey on Application and Future Prospects. *Environ. Technol. Innov.* **2019**, *13*, 264–276. [[CrossRef](#)]

9. Bandehali, S.; Miri, T.; Onyeaka, H.; Kumar, P. Current State of Indoor Air Phytoremediation Using Potted Plants and Green Walls. *Atmosphere* **2021**, *12*, 473. [[CrossRef](#)]
10. Zhang, R.; Yang, P.; Liu, S.; Wang, C.; Liu, J. Evaluation of the Methods for Estimating Leaf Chlorophyll Content with SPAD Chlorophyll Meters. *Remote Sens.* **2022**, *14*, 5144. [[CrossRef](#)]
11. Swoczyna, T.; Kalaji, H.M.; Bussotti, F.; Mojski, J.; Pollastrini, M. Environmental Stress—What Can We Learn from Chlorophyll a Fluorescence Analysis in Woody Plants? A Review. *Front. Plant Sci.* **2022**, *13*, 1048582. [[CrossRef](#)]
12. Boegh, E.; Houborg, R.; Bienkowski, J.; Braban, C.F.; Dalgaard, T.; van Dijk, N.; Dragosits, U.; Holmes, E.; Magliulo, V.; Schelde, K.; et al. Remote Sensing of LAI, Chlorophyll and Leaf Nitrogen Pools of Crop- and Grasslands in Five European Landscapes. *Biogeosciences* **2013**, *10*, 6279–6307. [[CrossRef](#)]
13. Uchino, H.; Watanabe, T.; Ramu, K.; Sahrawat, K.L.; Marimuthu, S.; Wani, S.P.; Ito, O. Calibrating Chlorophyll Meter (Spad-502) Reading by Specific Leaf Area for Estimating Leaf Nitrogen Concentration in Sweet Sorghum. *J. Plant Nutr.* **2013**, *36*, 1640–1646. [[CrossRef](#)]
14. Cerovic, Z.G.; Masdoumier, G.; Ghozlen, N.B.; Latouche, G. A New Optical Leaf-Clip Meter for Simultaneous Non-Destructive Assessment of Leaf Chlorophyll and Epidermal Flavonoids. *Physiol. Plant.* **2012**, *146*, 251–260. [[CrossRef](#)] [[PubMed](#)]
15. Shah, S.H.; Houborg, R.; McCabe, M.F. Response of Chlorophyll, Carotenoid and SPAD-502 Measurement to Salinity and Nutrient Stress in Wheat (*Triticum aestivum* L.). *Agronomy* **2017**, *7*, 61. [[CrossRef](#)]
16. Maier, P.M.; Keller, S.; Hinz, S. Deep Learning with WASI Simulation Data for Estimating Chlorophyll a Concentration of Inland Water Bodies. *Remote Sens.* **2021**, *13*, 718. [[CrossRef](#)]
17. Jin, D.; Lee, E.; Kwon, K.; Kim, T. A Deep Learning Model Using Satellite Ocean Color and Hydrodynamic Model to Estimate Chlorophyll-a Concentration. *Remote Sens.* **2021**, *13*, 2003. [[CrossRef](#)]
18. Aloisio, J.M.; Palmer, M.I.; Giampieri, M.A.; Tuininga, A.R.; Lewis, J.D. Spatially Dependent Biotic and Abiotic Factors Drive Survivorship and Physical Structure of Green Roof Vegetation. *Ecol. Appl.* **2017**, *27*, 297–308. [[CrossRef](#)]
19. Ojha, T.; Misra, S.; Raghuwanshi, N.S. Internet of Things for Agricultural Applications: The State of the Art. *IEEE Internet Things J.* **2021**, *8*, 10973–10997. [[CrossRef](#)]
20. Stevens, J.D.; Murray, D.; Diepeveen, D.; Toohey, D. Development and Testing of an IoT Spectroscopic Nutrient Monitoring System for Use in Micro Indoor Smart Hydroponics. *Horticulturae* **2023**, *9*, 185. [[CrossRef](#)]
21. Jurišić, M.; Radočaj, D.; Plaščak, I.; Galić Subašić, D.; Petrović, D. The Evaluation of the RGB and Multispectral Camera on the Unmanned Aerial Vehicle (UAV) for the Machine Learning Classification of Maize. *Poljoprivreda* **2022**, *28*, 74–80. [[CrossRef](#)]
22. Neupane, J.; Guo, W. Agronomic Basis and Strategies for Precision Water Management: A Review. *Agronomy* **2019**, *9*, 87. [[CrossRef](#)]
23. Purbajanti, E.D.; Slamet, W.; Fuskhah, E. Rosyida Effects of Organic and Inorganic Fertilizers on Growth, Activity of Nitrate Reductase and Chlorophyll Contents of Peanuts (*Arachis hypogaea* L.). *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *250*, 012048. [[CrossRef](#)]
24. 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](#)]
25. Ouafiq, E.M.; Saadane, R.; Chehri, A. Data Management and Integration of Low Power Consumption Embedded Devices IoT for Transforming Smart Agriculture into Actionable Knowledge. *Agriculture* **2022**, *12*, 329. [[CrossRef](#)]
26. Alshammari, T.; Ramadan, R.A.; Ahmad, A. Temporal Variations Dataset for Indoor Environmental Parameters in Northern Saudi Arabia. *Appl. Sci.* **2023**, *13*, 7326. [[CrossRef](#)]
27. Tan, L.; Zhou, L.; Zhao, N.; He, Y.; Qiu, Z. Development of a Low-Cost Portable Device for Pixel-Wise Leaf SPAD Estimation and Blade-Level SPAD Distribution Visualization Using Color Sensing. *Comput. Electron. Agric.* **2021**, *190*, 106487. [[CrossRef](#)]
28. Ma, D.; Carpenter, N.; Amatya, S.; Maki, H.; Wang, L.; Zhang, L.; Neeno, S.; Tuinstra, M.R.; Jin, J. Removal of Greenhouse Microclimate Heterogeneity with Conveyor System for Indoor Phenotyping. *Comput. Electron. Agric.* **2019**, *166*, 104979. [[CrossRef](#)]
29. Wu, Q.; Zhang, Y.; Zhao, Z.; Xie, M.; Hou, D. Estimation of Relative Chlorophyll Content in Spring Wheat Based on Multi-Temporal UAV Remote Sensing. *Agronomy* **2023**, *13*, 211. [[CrossRef](#)]
30. Sudu, B.; Rong, G.; Guga, S.; Li, K.; Zhi, F.; Guo, Y.; Zhang, J.; Bao, Y. Retrieving SPAD Values of Summer Maize Using UAV Hyperspectral Data Based on Multiple Machine Learning Algorithm. *Remote Sens.* **2022**, *14*, 5407. [[CrossRef](#)]
31. Degaspari, C.G.; Demétrio, C.A.; Jacob, J.F.d.O.; Ambrosano, G.B.; Rodrigues, P.H.V. Lantana Flowering at an Indoor Active Living Wall in a Light-Restricted Environment. *Ciênc. Rural.* **2023**, *53*, e20220493. [[CrossRef](#)]
32. Gräf, M.; Stangl, R.; Hood-Nowotny, R.; Kodym, A. Urban Farming in Indoor Settings: Nitrate Limits Compliance Check of Leafy Green Vegetables under LED Lighting. *Eur. J. Hortic. Sci.* **2020**, *85*, 321–328. [[CrossRef](#)]
33. Radočaj, D.; Šiljeg, A.; Plaščak, I.; Marić, I.; Jurišić, M. A Micro-Scale Approach for Cropland Suitability Assessment of Permanent Crops Using Machine Learning and a Low-Cost UAV. *Agronomy* **2023**, *13*, 362. [[CrossRef](#)]
34. Almalki, F.A.; Soufiene, B.O.; Alsamhi, S.H.; Sakli, H. A Low-Cost Platform for Environmental Smart Farming Monitoring System Based on IoT and UAVs. *Sustainability* **2021**, *13*, 5908. [[CrossRef](#)]
35. Morais, R.; Mendes, J.; Silva, R.; Silva, N.; Sousa, J.J.; Peres, E. A Versatile, Low-Power and Low-Cost IoT Device for Field Data Gathering in Precision Agriculture Practices. *Agriculture* **2021**, *11*, 619. [[CrossRef](#)]
36. Kitić, G.; Tagarakis, A.; Cselyuszka, N.; Panić, M.; Birgermajer, S.; Sakulski, D.; Matović, J. A New Low-Cost Portable Multispectral Optical Device for Precise Plant Status Assessment. *Comput. Electron. Agric.* **2019**, *162*, 300–308. [[CrossRef](#)]

37. Radočaj, D.; Šiljeg, A.; Marinović, R.; Jurišić, M. State of Major Vegetation Indices in Precision Agriculture Studies Indexed in Web of Science: A Review. *Agriculture* **2023**, *13*, 707. [[CrossRef](#)]
38. Girolimetto, D.; Venturini, V. Water Stress Estimation from NDVI-Ts Plot and the Wet Environment Evapotranspiration. *Adv. Remote Sens.* **2013**, *2*, 283–291. [[CrossRef](#)]
39. Agapiou, A.; Hadjimitsis, D.G.; Alexakis, D.D. Evaluation of Broadband and Narrowband Vegetation Indices for the Identification of Archaeological Crop Marks. *Remote Sens.* **2012**, *4*, 3892–3919. [[CrossRef](#)]
40. Wang, F.; Huang, J.; Tang, Y.; Wang, X. New Vegetation Index and Its Application in Estimating Leaf Area Index of Rice. *Rice Sci.* **2007**, *14*, 195–203. [[CrossRef](#)]
41. Barnes, E.M.; Clarke, T.R.; Richards, S.E.; Colaizzi, P.D.; Haberland, J.; Kostrzewski, M.; Waller, P.; Choi, C.; Riley, E.; Thompson, T.; et al. Coincident Detection of Crop Water Stress, Nitrogen Status and Canopy Density Using Ground-Based Multispectral Data. In Proceedings of the 5th International Conference on Precision Agriculture, Bloomington, MN, USA, 16–19 July 2000; pp. 1–15.
42. Jordan, C.F. Derivation of Leaf-Area Index from Quality of Light on the Forest Floor. *Ecology* **1969**, *50*, 663–666. [[CrossRef](#)]
43. Rorie, R.L.; Purcell, L.C.; Karcher, D.E.; King, C.A. The Assessment of Leaf Nitrogen in Corn from Digital Images. *Crop Sci.* **2011**, *51*, 2174–2180. [[CrossRef](#)]
44. Peñuelas, J.; Filella, I. Visible and Near-Infrared Reflectance Techniques for Diagnosing Plant Physiological Status. *Trends Plant Sci.* **1998**, *3*, 151–156. [[CrossRef](#)]
45. Gamon, J.A.; Surfus, J.S. Assessing Leaf Pigment Content and Activity with a Reflectometer. *New Phytol.* **1999**, *143*, 105–117. [[CrossRef](#)]
46. Chen, J.M. Evaluation of Vegetation Indices and a Modified Simple Ratio for Boreal Applications. *Can. J. Remote Sens.* **1996**, *22*, 229–242. [[CrossRef](#)]
47. Roujean, J.-L.; Breon, F.-M. Estimating PAR Absorbed by Vegetation from Bidirectional Reflectance Measurements. *Remote Sens. Environ.* **1995**, *51*, 375–384. [[CrossRef](#)]
48. Crippen, R.E. Calculating the Vegetation Index Faster. *Remote Sens. Environ.* **1990**, *34*, 71–73. [[CrossRef](#)]
49. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the Radiometric and Biophysical Performance of the MODIS Vegetation Indices. *Remote Sens. Environ.* **2002**, *83*, 195–213. [[CrossRef](#)]
50. Sripada, R.P.; Heiniger, R.W.; White, J.G.; Meijer, A.D. Aerial Color Infrared Photography for Determining Early In-Season Nitrogen Requirements in Corn. *Agron. J.* **2006**, *98*, 968–977. [[CrossRef](#)]
51. Bannari, A.; Asalhi, H.; Teillet, P.M. Transformed Difference Vegetation Index (TDVI) for Vegetation Cover Mapping. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Toronto, ON, Canada, 24–28 June 2002; Volume 5, pp. 3053–3055.
52. Gitelson, A.A. Wide Dynamic Range Vegetation Index for Remote Quantification of Biophysical Characteristics of Vegetation. *J. Plant Physiol.* **2004**, *161*, 165–173. [[CrossRef](#)]
53. Rios do Amaral, L.; Molin, J.P. The Effectiveness of Three Vegetation Indices Obtained from a Canopy Sensor in Identifying Sugarcane Response to Nitrogen. *Agron. J.* **2014**, *106*, 273–280. [[CrossRef](#)]
54. Salmerón Gómez, R.; Rodríguez Sánchez, A.; García, C.G.; García Pérez, J. The VIF and MSE in Ridge Regression. *Mathematics* **2020**, *8*, 605. [[CrossRef](#)]
55. Ullah, M.I.; Aslam, M.; Altaf, S.; Ahmed, M. Some New Diagnostics of Multicollinearity in Linear Regression Model. *Sains Malays.* **2019**, *48*, 2051–2060. [[CrossRef](#)]
56. Banerjee, C.; Mukherjee, T.; Pasillio, E. Feature Representations Using the Reflected Rectified Linear Unit (RRReLU) Activation. *Big Data Min. Anal.* **2020**, *3*, 102–120. [[CrossRef](#)]
57. Kaloev, M.; Krastev, G. Comparative Analysis of Activation Functions Used in the Hidden Layers of Deep Neural Networks. In Proceedings of the 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Ankara, Turkey, 11–13 June 2021; pp. 1–5.
58. Owoyemi, A.; Porat, R.; Lichter, A.; Doron-Faigenboim, A.; Jovani, O.; Koenigstein, N.; Salzer, Y. Evaluation of the Storage Performance of ‘Valencia’ Oranges and Generation of Shelf-Life Prediction Models. *Horticulturae* **2022**, *8*, 570. [[CrossRef](#)]
59. Zheng, K.; Bo, Y.; Bao, Y.; Zhu, X.; Wang, J.; Wang, Y. A Machine Learning Model for Photorespiration Response to Multi-Factors. *Horticulturae* **2021**, *7*, 207. [[CrossRef](#)]
60. Radočaj, D.; Jurišić, M.; Gašparović, M.; Plaščak, I.; Antonić, O. Cropland Suitability Assessment Using Satellite-Based Biophysical Vegetation Properties and Machine Learning. *Agronomy* **2021**, *11*, 1620. [[CrossRef](#)]
61. Huang, L.; Liu, Y.; Huang, W.; Dong, Y.; Ma, H.; Wu, K.; Guo, A. Combining Random Forest and XGBoost Methods in Detecting Early and Mid-Term Winter Wheat Stripe Rust Using Canopy Level Hyperspectral Measurements. *Agriculture* **2022**, *12*, 74. [[CrossRef](#)]
62. Radočaj, D.; Jurišić, M.; Tadić, V. The Effect of Bioclimatic Covariates on Ensemble Machine Learning Prediction of Total Soil Carbon in the Pannonian Biogeoregion. *Agronomy* **2023**, *13*, 2516. [[CrossRef](#)]
63. Belgiu, M.; Dragut, L. Random Forest in Remote Sensing: A Review of Applications and Future Directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [[CrossRef](#)]
64. Huber, F.; Yushchenko, A.; Stratmann, B.; Steinhage, V. Extreme Gradient Boosting for Yield Estimation Compared with Deep Learning Approaches. *Comput. Electron. Agric.* **2022**, *202*, 107346. [[CrossRef](#)]

65. Caturegli, L.; Gaetani, M.; Volterrani, M.; Magni, S.; Minelli, A.; Baldi, A.; Brandani, G.; Mancini, M.; Lenzi, A.; Orlandini, S.; et al. Normalized Difference Vegetation Index versus Dark Green Colour Index to Estimate Nitrogen Status on Bermudagrass Hybrid and Tall Fescue. *Int. J. Remote Sens.* **2020**, *41*, 455–470. [[CrossRef](#)]
66. Morlin Carneiro, F.; Angeli Furlani, C.E.; Zerbato, C.; Candida de Menezes, P.; da Silva Gírio, L.A.; Freire de Oliveira, M. Comparison between Vegetation Indices for Detecting Spatial and Temporal Variabilities in Soybean Crop Using Canopy Sensors. *Precis. Agric.* **2020**, *21*, 979–1007. [[CrossRef](#)]
67. Zeng, Y.; Hao, D.; Huete, A.; Dechant, B.; Berry, J.; Chen, J.M.; Joiner, J.; Frankenberg, C.; Bond-Lamberty, B.; Ryu, Y.; et al. Optical Vegetation Indices for Monitoring Terrestrial Ecosystems Globally. *Nat. Rev. Earth Environ.* **2022**, *3*, 477–493. [[CrossRef](#)]
68. Pradhan, A.M.S.; Kim, Y.-T. Rainfall-Induced Shallow Landslide Susceptibility Mapping at Two Adjacent Catchments Using Advanced Machine Learning Algorithms. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 569. [[CrossRef](#)]
69. McCaw, Z.R.; Colthurst, T.; Yun, T.; Furlotte, N.A.; Carroll, A.; Alipanahi, B.; McLean, C.Y.; Hormozdiari, F. DeepNull Models Non-Linear Covariate Effects to Improve Phenotypic Prediction and Association Power. *Nat. Commun.* **2022**, *13*, 241. [[CrossRef](#)] [[PubMed](#)]
70. Kigo, S.N.; Omondi, E.O.; Omolo, B.O. Assessing Predictive Performance of Supervised Machine Learning Algorithms for a Diamond Pricing Model. *Sci. Rep.* **2023**, *13*, 17315. [[CrossRef](#)] [[PubMed](#)]
71. Zhou, G.; Ni, Z.; Zhao, Y.; Luan, J. Identification of Bamboo Species Based on Extreme Gradient Boosting (XGBoost) Using Zhuhai-1 Orbita Hyperspectral Remote Sensing Imagery. *Sensors* **2022**, *22*, 5434. [[CrossRef](#)] [[PubMed](#)]
72. Radočaj, D.; Jug, I.; Vukadinović, V.; Jurišić, M.; Gašparović, M. The Effect of Soil Sampling Density and Spatial Autocorrelation on Interpolation Accuracy of Chemical Soil Properties in Arable Cropland. *Agronomy* **2021**, *11*, 2430. [[CrossRef](#)]
73. Graditi, G.; Ferlito, S.; Adinolfi, G. Comparison of Photovoltaic Plant Power Production Prediction Methods Using a Large Measured Dataset. *Renew. Energy* **2016**, *90*, 513–519. [[CrossRef](#)]
74. Ubbens, J.R.; Stavness, I. Deep Plant Phenomics: A Deep Learning Platform for Complex Plant Phenotyping Tasks. *Front. Plant Sci.* **2017**, *8*, 1190. [[CrossRef](#)]
75. Varga, I.; Radočaj, D.; Jurišić, M.; Markulj Kulundžić, A.; Antunović, M. Prediction of Sugar Beet Yield and Quality Parameters with Varying Nitrogen Fertilization Using Ensemble Decision Trees and Artificial Neural Networks. *Comput. Electron. Agric.* **2023**, *212*, 108076. [[CrossRef](#)]
76. Hosseini, M.; Powell, M.; Collins, J.; Callahan-Flintoft, C.; Jones, W.; Bowman, H.; Wyble, B. I Tried a Bunch of Things: The Dangers of Unexpected Overfitting in Classification of Brain Data. *Neurosci. Biobehav. Rev.* **2020**, *119*, 456–467. [[CrossRef](#)] [[PubMed](#)]
77. Gilbertson, J.K.; van Niekerk, A. Value of Dimensionality Reduction for Crop Differentiation with Multi-Temporal Imagery and Machine Learning. *Comput. Electron. Agric.* **2017**, *142*, 50–58. [[CrossRef](#)]
78. Sudhakar, P.; Latha, P.; Reddy, P.V. *Phenotyping Crop Plants for Physiological and Biochemical Traits*; Academic Press: Cambridge, MA, USA, 2016; ISBN 978-0-12-804110-9.
79. Giovos, R.; Tassopoulos, D.; Kalivas, D.; Lougkos, N.; Priovolou, A. Remote Sensing Vegetation Indices in Viticulture: A Critical Review. *Agriculture* **2021**, *11*, 457. [[CrossRef](#)]
80. Villoslada, M.; Bergamo, T.F.; Ward, R.D.; Burnside, N.G.; Joyce, C.B.; Bunce, R.G.H.; Sepp, K. Fine Scale Plant Community Assessment in Coastal Meadows Using UAV Based Multispectral Data. *Ecol. Indic.* **2020**, *111*, 105979. [[CrossRef](#)]
81. Naik, P.; Dalponte, M.; Bruzzone, L. Prediction of Forest Aboveground Biomass Using Multitemporal Multispectral Remote Sensing Data. *Remote Sens.* **2021**, *13*, 1282. [[CrossRef](#)]
82. Kaya, A.; Keceli, A.S.; Catal, C.; Yalic, H.Y.; Temucin, H.; Tekinerdogan, B. Analysis of Transfer Learning for Deep Neural Network Based Plant Classification Models. *Comput. Electron. Agric.* **2019**, *158*, 20–29. [[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.