



## Article

# Predicting Days to Maturity, Plant Height, and Grain Yield in Soybean: A Machine and Deep Learning Approach Using Multispectral Data

Paulo Eduardo Teodoro <sup>1</sup>, Larissa Pereira Ribeiro Teodoro <sup>1</sup>, Fábio Henrique Rojo Baio <sup>1</sup>, Carlos Antonio da Silva Junior <sup>2</sup>, Regimar Garcia dos Santos <sup>1</sup>, Ana Paula Marques Ramos <sup>3</sup>, Mayara Maezano Faita Pinheiro <sup>3</sup>, Lucas Prado Osco <sup>4</sup>, Wesley Nunes Gonçalves <sup>5</sup>, Alexandro Monteiro Carneiro <sup>6</sup>, José Marcato Junior <sup>5</sup>, Hemerson Pistori <sup>6,7</sup> and Luciano Shozo Shiratsuchi <sup>8,\*</sup>

<sup>1</sup> Department of Agronomy, Federal University of Mato Grosso do Sul, Rodovia MS 306, km. 305, Caixa Postal 112, Chapadão do Sul 79560-000, MS, Brazil; paulo.teodoro@ufms.br (P.E.T.); larissa\_ribeiro@ufms.br (L.P.R.T.); fabiobaio@ufms.br (F.H.R.B.); regimar.santos@ufms.br (R.G.d.S.)

<sup>2</sup> Department of Geography, State University of Mato Grosso, Av. dos Ingas, 3001-Jardim Imperial, Sinop 78555-000, MT, Brazil; carlosjr@unemat.br

<sup>3</sup> Post-Graduate Program of Environment and Regional Development, University of Western São Paulo, Rodovia Raposo Tavares, km 572–Limoeiro, Presidente Prudente 19067-175, SP, Brazil; anaramos@unoeste.br (A.P.M.R.); mayarafaita@gmail.com (M.M.F.P.)

<sup>4</sup> Faculty of Engineering and Architecture and Urbanism, University of Western São Paulo, Rodovia Raposo Tavares, km 572–Limoeiro, Presidente Prudente 19067-175, SP, Brazil; lucasosco@unoeste.br

<sup>5</sup> Faculty of Engineering, Architecture and Urbanism and Geography, Federal University of Mato Grosso do Sul, Cidade Universitária, Av. Costa e Silva–Pioneiros, Campo Grande 79070-900, MS, Brazil; wesley.goncalves@ufms.br (W.N.G.); jose.marcato@ufms.br (J.M.J.)

<sup>6</sup> Inovisão, Universidade Católica Dom Bosco, Av. Tamandaré, 6000, Campo Grande 79117-900, MS, Brazil; rf5974@ucdb.br (A.M.C.); pistori@ucdb.br (H.P.)

<sup>7</sup> Faculty of Computing, Federal University of Mato Grosso do Sul, Cidade Universitária, Av. Costa e Silva–Pioneiros, Campo Grande 79070-900, MS, Brazil

<sup>8</sup> AgCenter, School of Plant, Environmental and Soil Sciences, Louisiana State University, Baton Rouge, LA 70808, USA

\* Correspondence: LShiratsuchi@agcenter.lsu.edu



**Citation:** Teodoro, P.E.; Teodoro, L.P.R.; Baio, F.H.R.; da Silva Junior, C.A.; dos Santos, R.G.; Ramos, A.P.M.; Pinheiro, M.M.F.; Osco, L.P.; Gonçalves, W.N.; Carneiro, A.M.; et al. Predicting Days to Maturity, Plant Height, and Grain Yield in Soybean: A Machine and Deep Learning Approach Using Multispectral Data. *Remote Sens.* **2021**, *13*, 4632. <https://doi.org/10.3390/rs13224632>

Academic Editors: Yubin Lan, Wenjiang Huang, Kyeong-Hwan Lee, Yali Zhang and Muhammad Naveed Tahir

Received: 15 October 2021

Accepted: 15 November 2021

Published: 17 November 2021

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



**Copyright:** © 2021 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:** In soybean, there is a lack of research aiming to compare the performance of machine learning (ML) and deep learning (DL) methods to predict more than one agronomic variable, such as days to maturity (DM), plant height (PH), and grain yield (GY). As these variables are important to developing an overall precision farming model, we propose a machine learning approach to predict DM, PH, and GY for soybean cultivars based on multispectral bands. The field experiment considered 524 genotypes of soybeans in the 2017/2018 and 2018/2019 growing seasons and a multitemporal–multispectral dataset collected by embedded sensor in an unmanned aerial vehicle (UAV). We proposed a multilayer deep learning regression network, trained during 2000 epochs using an adaptive subgradient method, a random Gaussian initialization, and a 50% dropout in the first hidden layer for regularization. Three different scenarios, including only spectral bands, only vegetation indices, and spectral bands plus vegetation indices, were adopted to infer each variable (PH, DM, and GY). The DL model performance was compared against shallow learning methods such as random forest (RF), support vector machine (SVM), and linear regression (LR). The results indicate that our approach has the potential to predict soybean-related variables using multispectral bands only. Both DL and RF models presented a strong ( $r$  surpassing 0.77) prediction capacity for the PH variable, regardless of the adopted input variables group. Our results demonstrated that the DL model ( $r = 0.66$ ) was superior to predict DM when the input variable was the spectral bands. For GY, all machine learning models evaluated presented similar performance ( $r$  ranging from 0.42 to 0.44) for each tested scenario. In conclusion, this study demonstrated an efficient approach to a computational solution capable of predicting multiple important soybean crop variables based on remote sensing data. Future research could benefit from the information presented here and be implemented in subsequent processes related to soybean cultivars or other types of agronomic crops.

**Keywords:** precision agriculture; multispectral remote sensing data; shallow learner; deep neural network

## 1. Introduction

Farmers must consider crop genotype management and environmental factors (temperature, rainfall, soil, etc.) to obtain cultivars with maximum yield [1]. Recently, to help monitor such conditions, they can count on remote sensing systems and technologies to survey their production fields. Nonetheless, predicting agronomic variables such as days to maturity (DM), plant height (PH), and grain yield (GY), is still a challenging task, especially when considering indirect methods, such as multispectral data collected with remote sensing system based on unmanned aerial vehicle (UAV). Recently, the advent of new UAV platforms promoted rapid and high-detailed mapping of multiple farmlands, which generated high amounts of data to be evaluated and incorporated to support agricultural management [2–7].

In the last decade, remote sensing data has been processed with machine learning (ML) methods, which is a promising strategy to support varieties of evaluation and selection and other agricultural applications. In the related literature, it is possible to find a representative number of studies addressing crop yield prediction based on ML models for different crops, such as cherry tree (*Prunus avium* L.) [8], sugarcane (*Saccharum officinarum* L.) [9], wheat (*Triticum aestivum* L.) [10–12], potato (*Solanum tuberosum* L.) [10], coffee (*Coffea arabica* L.) [13], rice (*Oryza sativa* L.) [14], and maize (*Zea mays* L.) [15], among others. Review papers have [16,17] argued that the rapid technological advances in remote sensing systems and ML techniques can provide a cost-effective and efficient solution for better decision-making practices related to crop management in the next few years.

In the context mentioned above, it should be noted that agronomic variables inference for soybean (*Glycine max* (L.) Merrill) cultivars using only remote sensing data is still an unsolved and challenging task, mainly when considering multiple genotype varieties in different geographical areas and seasons. As agricultural activity represents an essential component of the economic sector for many countries, identifying or predicting variables such as DM, PH, and GY can support decision-making and benefit future crop management strategies [2,6,18–20]. Soybean is the world's major oilseed, accounting for 85% of the world's oil seeds, and given its economic importance, soybean cultivation area worldwide is steadily increasing [21]. Brazil, for example, is the world's largest producer of soybean, obtaining a record production estimated at 124.8 million tons in the last season, which represents an increase of 4.3% compared to the 2018/19 harvest [2]. The planted soybean area in Brazil grew 3% compared to the last season (2019/2020), from 35,874 to the current 36,949.8 hectares [22]. Thus, technological solutions involving intelligent methods and remote sensing data can improve yield and help to estimate other related agronomic variables while producing novel information to be incorporated to support crop management for subsequent seasons.

Previous studies demonstrated that other plant-related agronomic variables, such as biomass [3], chlorophyll content, macro and micronutrients [4], water content and stress effect [5], already returned satisfactory results using artificial intelligence techniques. However, predicting grain yield with remote sensing imagery is still somewhat difficult. Wei and Molin [23] explored relationships between soybean yield and plant components (number of grains—NG and thousand grains weight—TGW) using a regression approach. They demonstrated that TGW and NG presented, respectively, weak and strong linear relationships with yield. For maize yield prediction, Khanal et al. [24] implemented a combination of multiple agronomic and sensing variables into an ML environment and achieved satisfactory performances. These and other related research [25] demonstrated the importance of the use of artificial intelligence to model this dataset. These methods included most shallow learners, such as random forest, neural network, support vector machine,

and gradient boosting model. Nonetheless, in recent years, a more robust approach, known as deep learning (DL), has gained attention in this area of nontraditional statistical analysis.

DL is a computational intelligence technique based on artificial neural networks with multiple layers that adopt a hierarchical learning process [18,26]. These layers consist of activation, pooling, convolutional, recurrent, encode–decode, and others to form an architecture capable of learning from a given pattern [27]. These approaches often require an expressive amount of data compared to traditional or shallow methods to achieve high performance. Regardless, the learning capability of these networks is superior, mostly resulting in practical and reproducible alternatives; however, because of that, they also demand a high computational cost. Even so, in agriculture-related problems, it has demonstrated a high potential to model biomass phenotyping [28], counting plants and detecting plantation rows [29], and, related to this research, yield prediction [1,12,16,19,30–32]. While these studies have achieved considerable performance in estimating yield in different crops, there is still a lack of research comparing different machine learning methods to predict more than one variable, such as plant height (PH), days to maturity (DM), and grain yield (GY), which are considered important for an overall precision farming model.

As mentioned, in the precision agriculture context, several approaches have been implemented to obtain the maximum potential of machine learning models to predict agronomic variables. A concern regarding studies associating agronomic variables and multispectral imagery is that vegetation indices rely on a small number of available spectral bands and therefore do not utilize all the information conveyed by the plant genetic trait [16]. Thus, the question arises: Which spectral bands or vegetation indices (VIs) are better for the given task? Previous studies [3,5,15] evaluated the importance of different input variables configurations of remote sensing datasets using ML approaches related to plant analysis. Marques Ramos et al. [15] analyzed the individual contribution of VIs for a set of shallow learning models. There, the random forest (RF) returned the overall most effective estimative for maize yield. However, studies comparing different input variables configurations using both shallow and deep learning approaches are still lacking. As DL can extract key features from the data for estimation, amplifying aspects of the input that are important for prediction and suppressing irrelevant variations, it can be expected to have less dependency on the input data than shallow learning models [16,33]. Additionally, something that still needs to be further investigated is the performance evaluation of deep and shallow learning models to predict variables such as DM, PH, and GY, since they are important agronomic/plant parameters to support decision-making in soybean crops. To perform such evaluation, several soybean genotypes, considering a multitemporal dataset, should be required to produce representative conditions of farmers across the country.

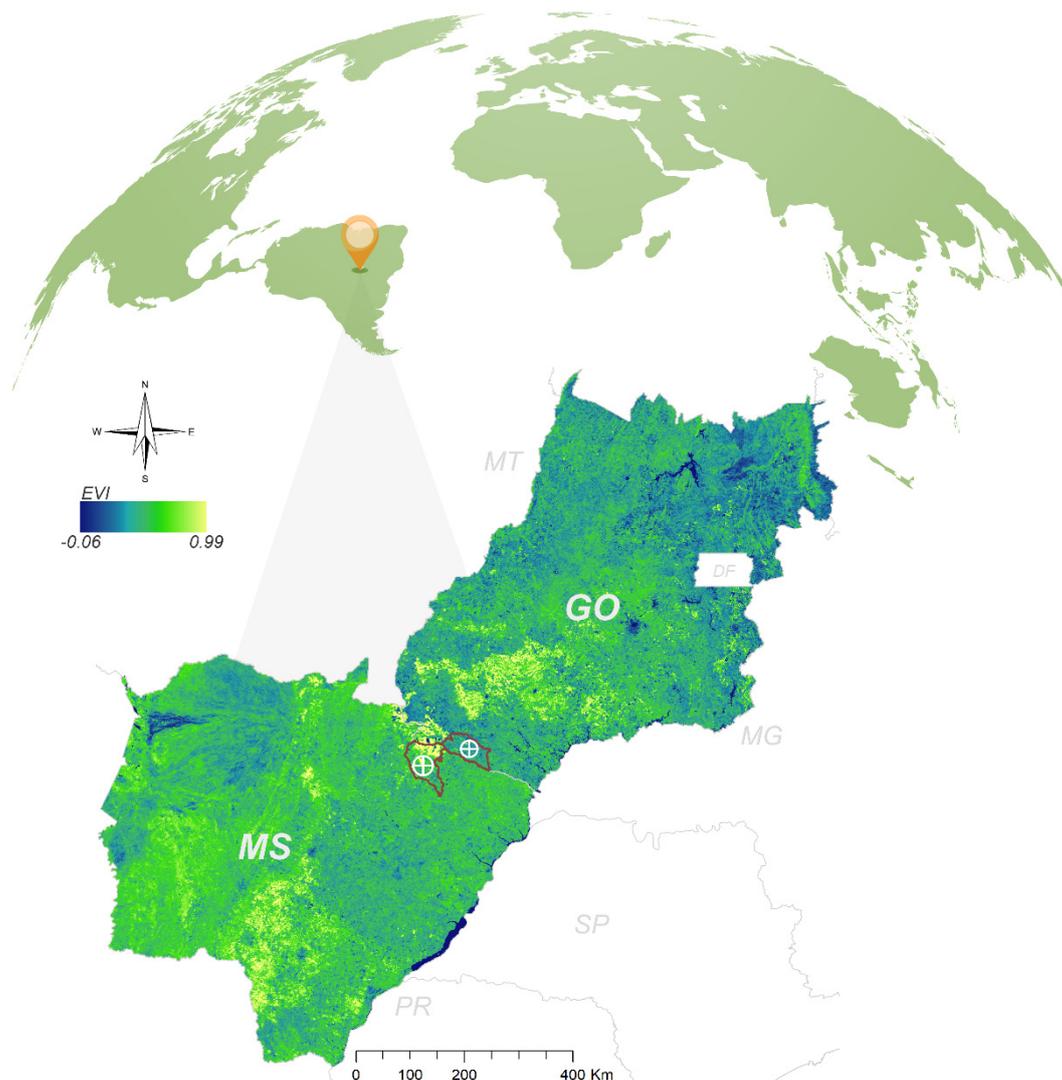
To help fulfil the aforementioned gap, here, we propose machine and deep learning approaches to predict soybean agronomic variables (DM, PH, and GY) using multispectral data from a UAV-based sensor. Our approach implemented a multilayer deep learning regression network, and we evaluated three different input configurations using spectral bands and VIs to infer these variables. For this, a total of 34 spectral indices and different combinations of spectral bands and VIs were used. Moreover, we compared the DL model against three shallow learning models: random forests (RF), support vector machine (SVM), and linear regression (LR), showing their potential to predict soybean-related variables. The main contribution of this study is to demonstrate a computational solution capable of predicting important soybean agronomic variables based on an efficient machine learning approach. In this regard, we consider a challenging situation: 524 genotypes of soybean in 2017/2018 and 2018/2019 crop seasons in four different geographical areas.

## 2. Materials and Methods

### 2.1. Field Trials

The data used in this research consist of experimental field trials conducted with soybean genotypes in the soybean macroregion 301, Brazil (Figure 1). The first trial was conducted in the 2017/2018 crop season in Chapadão do Sul/MS. The second and

third trials were conducted in Aporé/GO and Chapadão do Sul/MS, respectively, in the 2018/2019 crop season. The fourth trial was conducted in Chapadão do Sul in the 2019/2020 crop season. In all site-years, the crop seasons corresponded from October to January. We evaluated 123, 135, 103, and 163 soybean genotypes in trials 1, 2, 3, and 4, respectively, in a randomized block design with two replications. The experimental plots consisted of three 3 m rows spaced at 0.45 m. with a stand of 15 plants  $m^{-1}$ . The cultural treatments carried out during the tests followed the recommendation for the soybean crops in the Brazilian Cerrado environment.



**Figure 1.** Location of the studied areas in Chapadão do Sul (MS) and Aporé (GO), Brazil. EVI = enhanced vegetation index.

The variables evaluated in each trial were: plant height (PH, cm), days to maturity (DM), and grain yield (GY,  $kg\ ha^{-1}$ ). PH was obtained with the aid of a tape measure, measuring from the base of the plant to its apex. The DM was obtained by counting the number of days between emergence and maturation of more than 90% of the plants in the plot. Grain yield was obtained by harvesting 2 m long of the central row from each plot, weighing the grains, correcting for 13% moisture, and extrapolating to  $kg\ ha^{-1}$ .

## 2.2. Aerial Multispectral Image Acquisition and Vegetation Indices Calculation

At 60 days after crop emergence, the senseFly eBee<sup>®</sup> RTK fixed-wing UAV was used in all four trials. This equipment has autonomous take-off, flight plan, and landing con-

trol. The eBee was equipped with the senseFly Parrot Sequoia<sup>®</sup> multispectral sensor, a multispectral camera for agriculture that uses a sunshine sensor and an additional 16 MP RGB camera for scouting. The overflights were performed at 100 m altitude, allowing a spatial image resolution of 0.10 m. The overflights were carried out near the zenith due to the minimization of the shadows of the plants at 11 a.m., given that the multispectral sensor is a passive type (i.e., dependent on solar luminosity). Radiometric calibration was performed for the entire scene based on a calibrated reflective surface provided by the manufacturer. Multispectral reflectance images were obtained for green (550 nm ± 40 nm), red (660 nm ± 40 nm), red-edge (735 nm ± 10 nm), and near-infrared (NIR, 790 nm ± 40 nm) spectral bands. The information obtained at these wavelengths allows the calculation of various vegetation indices maps to be used in computational algorithms.

The images were mosaicked and orthorectified by the computer program Pix4Dmapper. The positional accuracy of the orthoimages was verified with ground control points (GPC) surveyed with real-time kinematic (RTK). Table 1 presents all the VIs used in experiments. The group of VIs included in this work is based on the Osco et al. [5] and Marques Ramos et al. [15] approaches, as they evaluated data from the same spectral regions.

**Table 1.** Vegetation indices used in the experiments.

Index	Equation
ARVI2 (Atmospherically Resistant Vegetation Index 2)	$-0.18 + 1.17 * ((R\lambda_{nir} - R\lambda_{red}) / (R\lambda_{nir} + R\lambda_{red}))$
ATSAVI (Adjusted Transformed soil-adjusted VI)	$1.22 * \left[ \frac{(R\lambda_{nir} - 1.22 * R\lambda_{red} - 0.03)}{(1.22 * R\lambda_{nir} + R\lambda_{red} - 1.22 * 0.03 + 0.08(1 + 1.22^2))} \right]$
BWDRVI (Blue-wide dynamic range vegetation index)	$0.1 * (R\lambda_{nir} + R\lambda_{red})$
CCCI (Canopy Chlorophyll Content Index)	$\frac{[(R\lambda_{nir} - R\lambda_{rededge})]}{[(R\lambda_{nir} + R\lambda_{rededge})] / \left[ \frac{(R\lambda_{nir} - R\lambda_{red})}{(R\lambda_{nir} + R\lambda_{red})} \right]}$
CIgreen (Chlorophyll Index Green)	$\frac{(R\lambda_{nir})}{R\lambda_{green}} - 1$
CIrededge (Chlorophyll Index RedEdge)	$\frac{(R\lambda_{nir})}{R\lambda_{rededge}} - 1$
CVI (Chlorophyll Vegetation Index)	$\frac{R\lambda_{nir} * (R\lambda_{red})}{R\lambda_{green}^2}$
DVI (Difference Vegetation Index)	$R\lambda_{nir} / R\lambda_{red}$
EVEI2 (Enhanced Vegetation Index 2)	$2.5 * (R\lambda_{nir} - R\lambda_{red}) / (R\lambda_{nir} + 2.4 * R\lambda_{red} + 1)$
GDVI (Difference NIR/Green Difference Vegetation Index)	$R\lambda_{nir} - R\lambda_{green}$
GEMI (Global Environment Monitoring Index)	$2 * (1 - 0.25 * 2) - ((R\lambda_{red} - 0.125) / (1 - R\lambda_{red}))$
GNDVI (Green Normalized Difference Vegetation Index)	$\frac{(R\lambda_{nir} - R\lambda_{green})}{(R\lambda_{nir} + R\lambda_{green})}$
GRNDVI (Green-Red NDVI)	$[R\lambda_{nir} - (R\lambda_{green} + R\lambda_{red})] / [R\lambda_{nir} + (R\lambda_{green} + R\lambda_{red})]$
GRVI (Green-Red Vegetation Index)	$\frac{(R\lambda_{green} - R\lambda_{red})}{(R\lambda_{green} + R\lambda_{red})}$
GSAVI (Green Soil Adjusted Vegetation Index)	$[(R\lambda_{nir} - R\lambda_{green}) / (R\lambda_{nir} + R\lambda_{green} + 0.5)] * 1.5$
GTVI (Green Triangle Vegetation Index)	$(NDVI + 0.5) / (NDVI + 0.5) * \left[ \left( \sqrt{NDVI} + 0.5 \right) \right]$
IPVI (Infrared Percentage Vegetation Index)	$R\lambda_{nir} / ((R\lambda_{nir} + R\lambda_{red}) / 2) * (NDVI + 1)$
LogR (Log Ratio)	$\text{Log}(R\lambda_{nir} / R\lambda_{red})$
MSAVI (Modified Soil Adjusted Vegetation Index)	$\left[ 2 * (R\lambda_{nir} + 1) - \sqrt{(2 * R\lambda_{nir} + 1)^2 - 8 * (R\lambda_{nir} - R\lambda_{red})} \right] / 2$
MSRNir_Red (Modified Simple Ratio NIR/RED)	$\frac{(R\lambda_{nir} / R\lambda_{red} - 1)}{\sqrt{(R\lambda_{nir} / R\lambda_{red} + 1)}}$
NDRE (Normalized Difference Red-Edge Index)	$\frac{R\lambda_{nir} - R\lambda_{rededge}}{\lambda_{nir} + R\lambda_{rededge}}$
NDVI (Normalized Difference Vegetation Index)	$\frac{(R\lambda_{nir} - R\lambda_{red})}{(R\lambda_{nir} + R\lambda_{red})}$

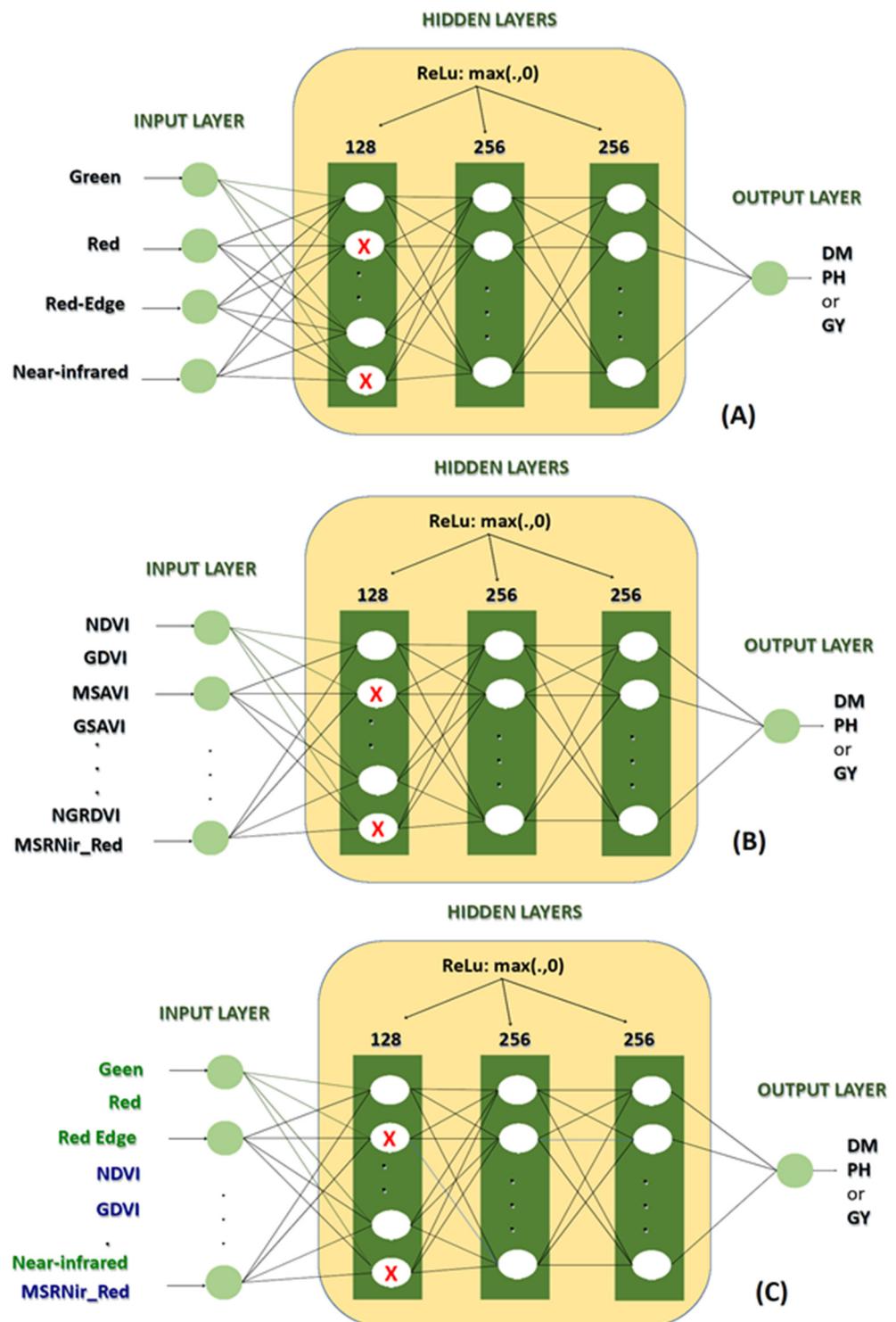
Table 1. Cont.

Index	Equation
NGRDI (Normalized Green-Red Difference Index)	$\frac{(R\lambda_{green} - R\lambda_{red})}{(R\lambda_{green} + R\lambda_{red})}$
NormR1 (Normalized G)	$\frac{R\lambda_{green}}{(R\lambda_{nir} + R\lambda_{red} + R\lambda_{green})}$
NormR2 (Normalized NIR)	$\frac{R\lambda_{nir}}{(R\lambda_{nir} + R\lambda_{red} + R\lambda_{green})}$
NormR3 (Normalized R)	$\frac{R\lambda_{red}}{(R\lambda_{nir} + R\lambda_{red} + R\lambda_{green})}$
RGR (Red Green Ratio Index)	$\frac{(R\lambda_{red})}{R\lambda_{green}}$
RI (Redness Index)	$R\lambda_{red} - R\lambda_{green} / R\lambda_{red} + R\lambda_{green}$
RRI 1	$\frac{R\lambda_{nir}}{R\lambda_{rededge}}$
SRQT_IR_R (Square root of the NIR/Red ratio)	$\sqrt{R\lambda_{nir} / R\lambda_{red}}$
SRRed_NIR	$\frac{R\lambda_{red}}{R\lambda_{nir}}$
TNDVI (Transformed NDVI)	$\sqrt{[(R\lambda_{nir} - R\lambda_{red}) / (R\lambda_{nir} + R\lambda_{red}) + 0.5]}$
TVI (Transformed Vegetation Index)	$\sqrt{[NDVI + 0.5]}$
WDRVI (Wide Dynamic Range Vegetation Index)	$(0.1 * ((R\lambda_{nir} - R\lambda_{red})) / (0.1 * (R\lambda_{nir} + R\lambda_{red})))$

NIR = near-infrared;  $R\lambda$  = wavelength reflectance; RRI = Simple Ratio NIR/Rededge RedEdge Ratio Index 1; SRQT\_IR\_R = square root of the NIR/Red ratio; SRRed\_NIR = Simple Ratio Red/NIR Ratio Vegetation-Index.

### 2.3. Machine Learning

A deep learning (DL) model was implemented and tested against three shallow learning models: random forests (RF), support vector machine (SVM), and linear regression (LR). For all models, three data input configurations were adopted: (1) spectral wavelengths (WL), (2) spectral vegetation indices (VI), and (3) a combination between spectral wavelengths and vegetation indices (WLVI). The DL architecture proposed in this work (Figure 2) comprises three dense (completely connected) hidden layers, with 128, 256, and 256 neurons, respectively. The first hidden layer used dropout at 50% as a regularization strategy. An adaptive subgradient method, adagrad, was used for training the model over 2000 epochs, with a learning rate of 0.01, batch size of 32, and an early stop with 400 epochs (patience). The activation function used for all neurons was a rectified linear unit (ReLU), and the weights were randomly initialized using a standard Gaussian method. The mean squared error (MSE) was used as the loss function. The number of neurons in the input layer depended on the three different input configurations used in the experiments, as follows: (1) WL—4 neurons, (2) VI—34 neurons, and (3) WLVI—38 neurons. The output layer comprises only one neuron, as the network was trained for a regression approach. The number of trainable parameters (network weights) varied from 99,713 (WL) to 104,065 (WLVI).



**Figure 2.** Architecture of the three deep learning networks (DL) proposed in the study. (A) corresponds to the network using the spectral bands of the images as input layer; (B) corresponds to the network using the vegetation indices as input layer; (C) corresponds to the network using the spectral bands and vegetation indices as input layer.

The RF algorithm was trained using 100% of the training set as the bag size and 100 trees. SVM used the optimizer proposed by Shevade et al. [29] that improves sequential minimal optimization (SMO) for regression with a polynomial kernel. The LR performs a ridge regression using the Akaike information criterion (ACI) for model selection. The

ridge parameter was set to  $1 \times 10^{-8}$ . All four models (DL, RF, SVM, and LR) were trained and tested in 30 repetitions, using a random split of 80% for training and 20% for testing each time. The deep learning model further used 20% of the training set as validation. The other models (RF, SVM, and LR) did not need a validation set. For all the three input configurations, WL, VI, and WLVI, the models were trained to independently predict PH, DM, and GY output variables, given 36 tests.

#### 2.4. Statistical Analyses

Mean absolute error (MAE), root mean squared error (RMSE), and Pearson's correlation coefficient ( $r$ ) between the observed and estimated values were obtained in each test and repetition. Boxplots were created for all these 36 test results. A two-way ANOVA was performed for each output variable, PH, DM, and GY, followed by the Scott–Knott test at a 5% significance to identify the input configurations and ML models that provided the highest  $r$  means. In this ANOVA, the first factor consisted of the tested models (DL, RF, SVM, and LR), while the second factor was composed of the data input configurations (WL, VI, and WLVI).

### 3. Results

The results indicated that there was a significant interaction between the ML models (DL, RF, SVM, and LR) and the input settings (WL, VI, and WLVI) for the mean absolute error (MAE), root mean squared error (RMSE), and Pearson's correlation coefficient ( $r$ ) between observed and estimated values for DM (Table 2) and PH (Table 3). The results of the significant interaction between models versus input configurations for PH and DM demonstrate that there is a variable relationship between these factors, i.e., the best model depends on the configuration used, and vice versa. For GY, there were significant differences between the ML models tested for MAE and RMSE (Table 4). The  $r$  was statistically equal for ML models and tested inputs for GY.

**Table 2.** Grouping of means of mean absolute error (MAE), root mean squared error (RMSE), and Pearson's correlation coefficient ( $r$ ) between estimated and observed values for days to maturity in soybean obtained with different machine learning models and input configurations.

Model	Input		
	WL	VI	WLVI
	MAE		
DL	6.05 Bc	7.45 Aa	7.62 Aa
RF	6.24 Ac	6.09 Ac	6.04 Ac
SVM	7.11 Ab	6.65 Bb	6.38 Bb
LR	7.37 Aa	7.43 Aa	7.41 Aa
	RMSE		
DL	8.01 Cd	10.23 Ba	10.58 Aa
RF	8.58 Ac	8.28 Ac	8.23 Ad
SVM	9.76 Aa	9.25 Bb	8.82 Cc
LR	9.34 Ab	9.41 Ab	9.39 Ab
	$r$		
DL	0.66 Aa	0.57 Bb	0.54 Cc
RF	0.62 Ab	0.65 Aa	0.65 Aa
SVM	0.46 Cd	0.53 Bc	0.58 Ab
LR	0.51 Ac	0.50 Ad	0.50 Ad

Means followed by the same uppercase letters in the same row and the same lowercase letters in the same column do not differ by the Scott–Knott test at 5% probability. DL: deep learning (DL); RF: random forests; SVM: support vector machine; LR: linear regression; WL: spectral bands wavelengths; VI: spectral vegetation indices; WLVI: a combination between spectral bands and vegetation indices.

**Table 3.** Grouping of means of mean absolute error (MAE), root mean squared error (RMSE), and Pearson’s correlation coefficient (r) between estimated and observed values for plant height in soybean obtained with different machine learning models and input configurations.

Model	Input		
	WL	VI	WLVI
	MAE		
DL	8.32 Cb	10.89 Ba	11.92 Aa
RF	8.38 Ab	8.09 Ac	8.11 Ac
SVM	8.98 Aa	8.55 Bb	8.49 Bb
LR	9.03 Aa	8.65 Bb	8.67 Bb
	RMSE		
DL	10.51 Cb	13.38 Ba	14.55 Aa
RF	10.88 Ab	10.49 Ac	10.51 Ac
SVM	11.77 Aa	11.05 Bb	10.97 Bb
LR	11.78 Aa	11.24 Bb	11.21 Bb
	r		
DL	0.79 Aa	0.77 Ab	0.75 Bb
RF	0.77 Ba	0.79 Aa	0.79 Aa
SVM	0.73 Bb	0.76 Ac	0.77 Ab
LR	0.73 Bb	0.75 Ac	0.75 Ab

Means followed by the same uppercase letters in the same row and the same lowercase letters in the same column do not differ by the Scott–Knott test at 5% probability. DL: deep learning (DL); RF: random forests; SVM: support vector machine; LR: linear regression; WL: spectral bands wavelengths; VI: spectral vegetation indices; WLVI: a combination between spectral bands and vegetation indices.

**Table 4.** Grouping of means of mean absolute error (MAE), root mean squared error (RMSE), and Pearson’s correlation coefficient (r) between estimated and observed values for grain yield in soybean obtained with different machine learning models.

Model	MAE	RMSE	r
DL	788.31 b	1000.48 b	0.45 a
RF	807.16 a	1025.59 a	0.42 a
SVM	787.87 b	1010.11 b	0.44 a
LR	790.88 b	105.06 b	0.43 a

Means followed by the same uppercase letters in the same row and the same lowercase letters in the same column do not differ by the Scott–Knott test at 5% probability. DL: deep learning (DL); RF: random forests; SVM: support vector machine; LR: linear regression.

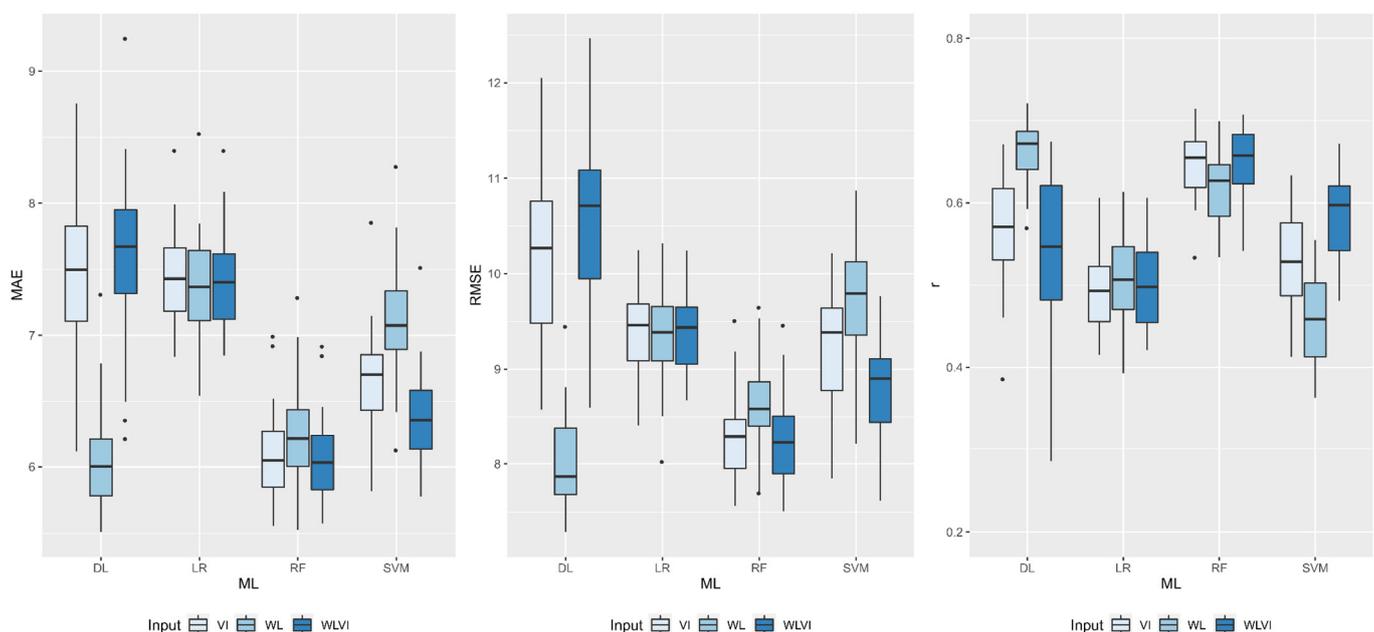
Table 2 demonstrates the unfolding of the significant interaction between ML and input configuration for MAE, RMSE, and r between estimated and observed values for DM in soybean. When analyzing the input for DL, we observed that WL had the lowest means for MAE (6.05) and RMSE (8.01) and the highest mean for r (0.66). For RF and LR, there were no statistical differences between the tested inputs. For SVM, the input with WLVI provided statistically lower MAE (6.38) and RMSE (8.82) values and higher r (0.58). When comparing the ML models using the WL as input, we found that DL had the lowest means for MAE (6.05) and RMSE (8.01), as well as the highest mean for r (0.66). However, when using VI or WLVI as inputs, RF was the best model due to the lower MAE (6.65 and 6.38, respectively) and RMSE (8.28 and 8.23, respectively) and the higher r (0.65 for both inputs).

The unfolding of the significant interaction between ML and input for MAE, RMSE, and between estimated and observed values for plant height in soybean is shown in Table 3. When analyzing the input for DL, we observed that WL had the lowest means for MAE

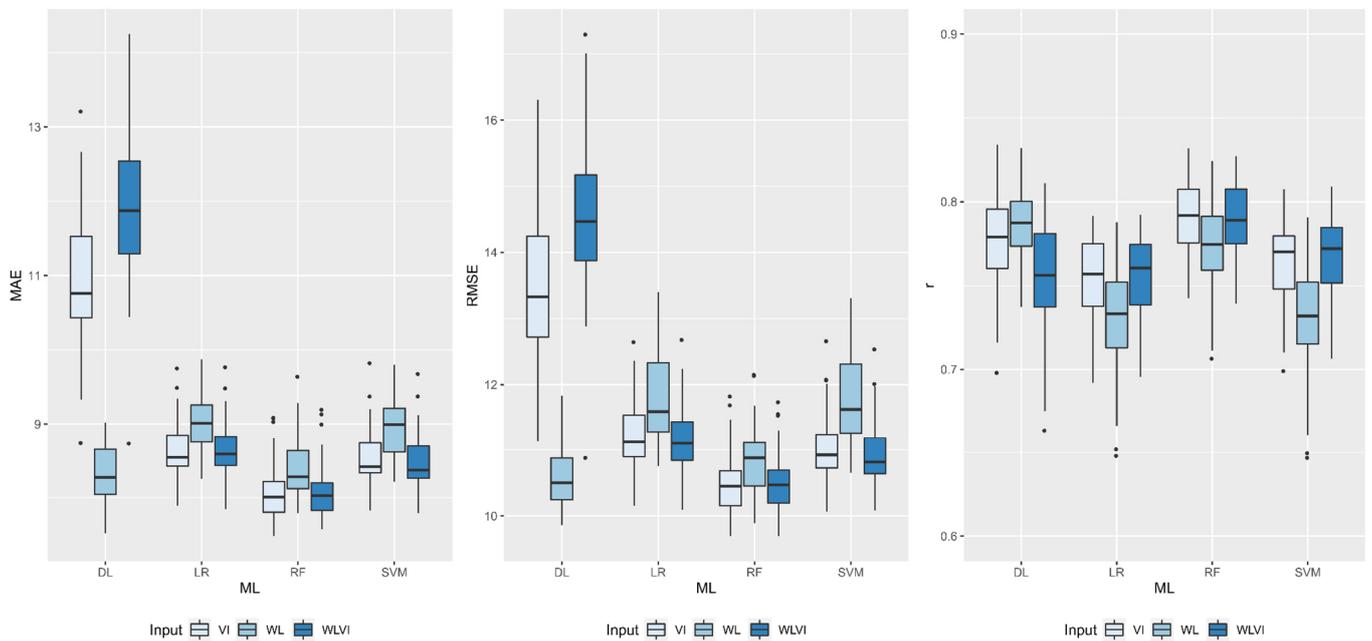
and RMSE and the highest mean for  $r$ . For RF and LR, there were no statistical differences between the tested inputs for MAE and RMSE, but for  $r$ , the best inputs were VI and WLVI. For SVM and LR, the inputs with VI and WLVI provided statistically lower MAE and RMSE values and higher  $r$ . When comparing the ML models using the WL as input, we found that DL and RF had the lowest means of MAE and RMSE, in addition to the highest means of  $r$ . When using VI or WLVI as inputs, RF was the best model due to the lower MAE and RMSE means and the highest  $r$  mean.

For GY, there was no significant interaction between inputs and ML models. The tested inputs did not differ statistically from each other. There were differences only between ML models in isolation for MAE and RMSE (Table 4). Among these, the DL, SVM, and LR models stood out for presenting lower MAE and RMSE in relation to the RF. For  $r$ , there was no difference between the ML models.

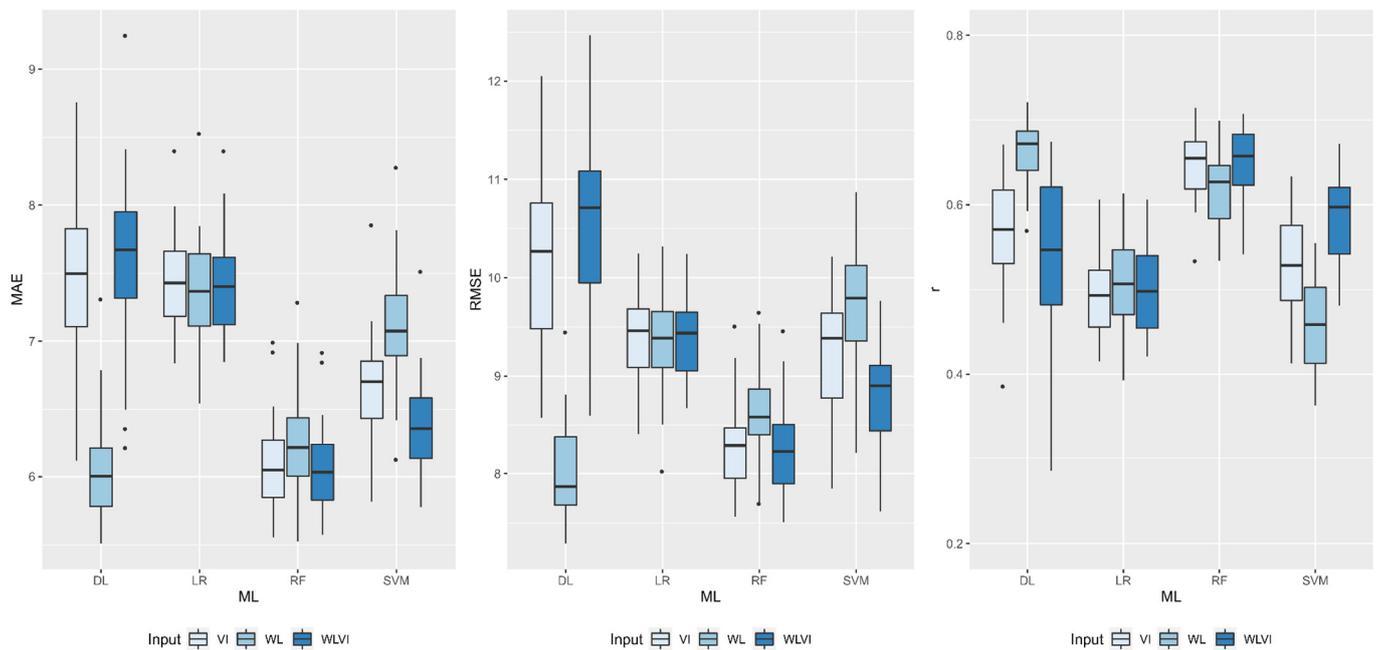
The boxplots for mean absolute error (MAE), root mean squared error (RMSE), and Pearson's correlation coefficient ( $r$ ) values between estimated and observed values for days to maturity (DM), plant height (PH), and grain yield (GY) in soybean obtained with different machine learning models (DL, RF, SVM, and LR) and input configurations (WL, VI, and WLVI) are grouped in Figures 3–5, respectively. In general, the combination between DL and WL provided greater accuracy for all variables evaluated, as it obtained lower median values for MAE and RMSE and higher values for  $r$ . Furthermore, in this combination, the boxes were smaller, which indicates less dispersion between the results of the 30 folds used to obtain the MAE, RMSE, and  $r$ . RF had slightly lower median accuracy values than DL when using WL as input. However, for the other inputs (VI and WLVI), this model had the lowest median values for MAE and RMSE, in addition to the highest values for  $r$ . However, it is important to highlight that the boxes for this model showed higher variability.



**Figure 3.** Boxplot for mean absolute error (MAE), root mean squared error (RMSE), and Pearson's correlation coefficient ( $r$ ) between estimated and observed values for days to maturity in soybean obtained with different machine learning models (DL, RF, SVM, and LR) and input configurations. DL: deep learning (DL); RF: random forests; SVM: support vector machine; LR: linear regression; WL: spectral bands wavelengths; VI: spectral vegetation indices; WLVI: a combination between spectral bands and vegetation indices.



**Figure 4.** Boxplot for mean absolute error (MAE), root mean squared error (RMSE), and Pearson's correlation coefficient ( $r$ ) between estimated and observed values for plant height in soybean obtained with different machine learning models (DL, RF, SVM, and LR) and input configurations. DL: deep learning (DL); RF: random forests; SVM: support vector machine; LR: linear regression; WL: spectral bands wavelengths; VI: spectral vegetation indices; WLVI: a combination between spectral bands and vegetation indices.



**Figure 5.** Boxplot for mean absolute error (MAE), root mean squared error (RMSE), and Pearson's correlation coefficient ( $r$ ) between estimated and observed values for grain yield in soybean obtained with different machine learning models (DL, RF, SVM, and LR) and input configurations. DL: deep learning (DL); RF: random forests; SVM: support vector machine; LR: linear regression; WL: spectral bands wavelengths; VI: spectral vegetation indices; WLVI: a combination between spectral bands and vegetation indices.

#### 4. Discussion

The significant interaction between inputs and ML models for DM and PH indicates differences between the models when using WL, VI, or WLVI as inputs. In the other sense, there are differences between the inputs for each ML model tested. The results obtained from MAE, RMSE, and  $r$  between estimated and observed values for DM and PH variables indicated a slightly superior performance of the proposed DL approach compared to the other traditional machine learning methods when only the WL are considered. DL methods extract features automatically, and this provides an advantageous approach. As verified in our experiments, the establishment of vegetation indices was not required, which implies an easy alternative to soybean variable prediction task. Another important observation regarding DL methods is the ability of updating novel training information based on a transfer learning strategy. Even though DL methods are known for demanding a computational power higher than ML approaches, the mentioned ability should be considered, especially in precision-farming tasks, as deep networks are more adaptive to new areas, images, and variables than shallow methods.

When considering only vegetation indices (VI) or vegetation indices combined with the wavelengths (WLVI) for both DM and PH, we found that RF outperformed the other models (lowest means of MAE and RMSE and highest means of  $r$ ). The highest performance found here using RF possibly occurred due to the internal structure of the algorithm, which is based on multiple decision tree sets. Recent studies have classified RF as an effective and versatile machine learning method for crop yield predictions [4,10,15]. Moreover, RF has been considered superior to other shallow learning algorithms because it can easily handle many model parameters, reduce estimate bias, and has no problems with overfitting. These findings are similar to previous reports on the prediction of agronomic traits using spectral data and machine learning. Yu et al. [20] found a good accuracy in yield prediction ( $r = 0.82$ ) of soybean breeding lines by using the RF model on data obtained with the UAV-based high-throughput phenotyping (HTP) platform. Sakamoto [18], when carrying out a MODIS-based crop yield estimation for corn and soybean with the RF algorithm, verified that the method provided an accurate estimate of yield reduction in soybeans caused by a drought that occurred during late vegetative stages. However, as performed here, the RF and shallow models have been little studied in predicting the maturity and plant height in soybean.

Days to maturity (DM) has been an important trait in studies aimed at improving soybean crops since, through this trait, it is possible to measure the earliness of cultivars. Earliness is one of the main targets currently required by the Brazilian soybean market. This occurs because farmers of the major producing regions, located in the Cerrado biome, can grow corn or cotton in the second season, between February and July, increasing the profitability of the agricultural system. Furthermore, early-cycle genotypes remain in the field for less time and are subject to less disease pressure during the late growing season, which causes considerable losses in crop yield [34]. However, measuring earliness is a time-demanding and labor-intensive task [6,35], since it requires counting, in the field, the number of days until each genotype blooms or matures. In this aspect, the approach proposed here may help to minimize this cost by implementing only remote sensing and computational options to perform said task.

Likewise, plant height (PH) is another important feature for improved soybean cultivars, since plants must be up to 1 m in height to avoid lodging and losses in mechanized harvesting. Several studies on genetic progress and breeding in soybean have identified a decrease in plant height and consequent reduction in the lodging score in modern cultivars [36–38]. However, PH assessment is also a time- and labor-consuming task. Thus, using a machine learning method that only needs spectral bands or vegetation indices from UAV-based imagery, agricultural practices can benefit from this approach to estimate PH with satisfactory performance. In a previous study, PH was also predicted with machine learning methods and obtained good accuracies, as in here, but in maize plants [7].

UAV-based multispectral imagery arises as a fast, cost-effective, and reliable technique for assessing crop attributes of interest, which is critical for high-throughput phenotyping in agriculture [39]. The literature has already reported the prediction of important agronomic variables (e.g., plant height, biomass, and grain yield) by using remote sensing techniques [2,3,7,15,20,40]. Nonetheless, studies aiming to predict crop maturity are recent, especially in soybean, and research using shallow and deep machine learning approaches is still scarce in the literature [20]. To the best of our knowledge, this is one of the first studies to perform a prediction for soybean maturity and plant height with shallow/deep learning-based methods applied to multispectral data considering multiple agronomic variables (PH, GY, and DM) at different cultivars, sites, and crop years. Our findings show that it is possible to predict days to maturity and plant height in soybean. The use of artificial intelligence techniques associated with data obtained by remote sensing shows to be a fast, low-cost, and efficient approach to be adopted in precision-farming practices, supporting crop management and helping in the decision-making process.

In this study, our interest was to investigate the performance of different machine learning methods and input configurations to predict days for maturation, plant height, and grain yield for soybeans using only multispectral UAV images. Overall, RF outperformed the other models when considering as inputs all of the VI and WLVI variables. This provided better accuracy in both the DM and PH predictions than when considering only the spectral bands. Yet, when using only the spectral bands as input, the DL model returned the overall best performance. We verified that the DL model can predict DM and PH in soybean using lesser variables, which is an advantageous and promising approach in terms of implementation and spectral data analysis.

The approach adopted here can be implemented with different datasets on different soybean cultivars. For further studies, we recommend that this approach be tested for different crop traits and using a larger number of wavelengths as input in the models.

## 5. Conclusions

We mainly verified significant interaction between machine learning models and input settings for MAE, RMSE, and  $r$  between observed and estimated values for both DM and PH traits. Our findings showed that the DL model outperformed the other models when using only WL as input configuration. The RF algorithm performed better when considering the VI, both single and combined with spectral bands (WLVI). Regardless, mainly because of its characteristics, the DL approach may learn more patterns in new datasets, which is an advantage over the traditional shallow learners. We conclude that our method can be adopted in future research to evaluate other traits in soybean. With this, we observed that this procedure can support the decision-making process on variety selection and yield prediction and supply more information to assist precision agriculture practices.

**Author Contributions:** P.E.T., L.P.R.T., F.H.R.B. and C.A.d.S.J. designed the field trials and collected the phenotypic data. H.P. performed all statistical analyses. P.E.T., L.P.R.T., A.P.M.R., L.P.O. and J.M.J. designed the manuscript, and produced a draft of the manuscript. R.G.d.S., M.M.F.P., W.N.G., A.M.C. and L.S.S. contributed with a critical review of the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Council for Scientific and Technological Development (CNPq), grant number 303767/2020-0, 303559/2019-5, 433783/2018-4, 314902/2018-0, and 304173/2016-9; Coordenação de Aperfeiçoamento de Pessoal de Nível Superior–Brazil (CAPES), grant number 88881.311850/2018-01, Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul (FUNDECT), grant number 71/019.039/2021, 59/300.066/2015, 23/200.248/2014, and 59/300.095/2015.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

**Acknowledgments:** The authors acknowledge the support of UFMS (Federal University of Mato Grosso do Sul), UCDB (Dom Bosco Catholic University), CNPq (National Council for Scientific and Technological), CAPES (Coordination for the Improvement of Higher Education Personnel–Finance code 001), and NVIDIA Corporation.

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

## References

1. Khaki, S.; Wang, L.; Archontoulis, S.V. A CNN-RNN Framework for Crop Yield Prediction. *Front. Plant Sci.* **2020**, *10*, 1750. [[CrossRef](#)]
2. da Silva, E.E.; Baio, F.H.R.; Teodoro, L.P.R.; Junior, C.A.D.S.; Borges, R.S.; Teodoro, P.E. UAV-multispectral and vegetation indices in soybean grain yield prediction based on in situ observation. *Remote Sens. Appl. Soc. Environ.* **2020**, *18*, 100318. [[CrossRef](#)]
3. 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](#)]
4. Osco, L.; Junior, J.; Ramos, A.; Furuya, D.; Santana, D.; Teodoro, L.; Gonçalves, W.; Baio, F.; Pistori, H.; Junior, C.; et al. Leaf Nitrogen Concentration and Plant Height Prediction for Maize Using UAV-Based Multispectral Imagery and Machine Learning Techniques. *Remote Sens.* **2020**, *12*, 3237. [[CrossRef](#)]
5. Osco, L.P.; Ramos, A.P.M.; Moriya, É.A.S.; Bavaresco, L.G.; De Lima, B.C.; Estrabis, N.; Pereira, D.R.; Creste, J.E.; Júnior, J.M.; Gonçalves, W.N.; et al. Modeling Hyperspectral Response of Water-Stress Induced Lettuce Plants Using Artificial Neural Networks. *Remote Sens.* **2019**, *11*, 2797. [[CrossRef](#)]
6. Zhou, J.; Yungbluth, D.C.; Vong, C.N.; Scaboo, A.M.; Zhou, J. Estimation of maturity date of soybean breeding lines using UAV-based imagery. *Remote Sens.* **2019**, *11*, 2075. [[CrossRef](#)]
7. Osco, L.P.; Ramos, A.P.M.; Pinheiro, M.M.F.; Moriya, A.S.; Imai, N.N.; Estrabis, N.; Ianczyk, F.; De Araújo, F.F.; Liesenberg, V.; Jorge, L.A.D.C.; et al. A Machine Learning Framework to Predict Nutrient Content in Valencia-Orange Leaf Hyperspectral Measurements. *Remote Sens.* **2020**, *12*, 906. [[CrossRef](#)]
8. Amatya, S.; Karkee, M.; Gongal, A.; Zhang, Q.; Whiting, M.D. Detection of cherry tree branches with full foliage in planar architecture for automated sweet-cherry harvesting. *Biosyst. Eng.* **2015**, *146*, 3–15. [[CrossRef](#)]
9. Vani, S.; Sukumaran, R.; Savithri, S. Prediction of sugar yields during hydrolysis of lignocellulosic biomass using artificial neural network modeling. *Bioresour. Technol.* **2015**, *188*, 128–135. [[CrossRef](#)]
10. Jeong, J.H.; Resop, J.; Mueller, N.D.; Fleisher, D.; Yun, K.; Butler, E.E.; Timlin, D.; Shim, K.-M.; Gerber, J.; Reddy, V.R.; et al. Random Forests for Global and Regional Crop Yield Predictions. *PLoS ONE* **2016**, *11*, e0156571. [[CrossRef](#)]
11. Pantazi, X.; Moshou, D.; Alexandridis, T.; Whetton, R.; Mouazen, A. Wheat yield prediction using machine learning and advanced sensing techniques. *Comput. Electron. Agric.* **2016**, *121*, 57–65. [[CrossRef](#)]
12. Hunt, M.L.; Blackburn, G.A.; Carrasco, L.; Redhead, J.W.; Rowland, C.S. High resolution wheat yield mapping using Sentinel-2. *Remote Sens. Environ.* **2019**, *233*, 111410. [[CrossRef](#)]
13. Ramos, P.; Prieto, F.; Montoya, E.; Oliveros, C. Automatic fruit count on coffee branches using computer vision. *Comput. Electron. Agric.* **2017**, *137*, 9–22. [[CrossRef](#)]
14. Su, Y.-X.; Xu, H.; Yan, L.-J. Support vector machine-based open crop model (SBOCM): Case of rice production in China. *Saudi J. Biol. Sci.* **2017**, *24*, 537–547. [[CrossRef](#)]
15. 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](#)]
16. Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine Learning Approaches for Crop Yield Prediction and Nitrogen Status Estimation in Precision Agriculture: A Review. *Comput. Electron. Agric.* **2018**, *151*, 61–69. [[CrossRef](#)]
17. Liakos, K.G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D. Machine learning in agriculture: A review. *Sensors* **2018**, *18*, 2674. [[CrossRef](#)] [[PubMed](#)]
18. Sakamoto, T. Incorporating environmental variables into a MODIS-based crop yield estimation method for United States corn and soybeans through the use of a random forest regression algorithm. *ISPRS J. Photogramm. Remote Sens.* **2019**, *160*, 208–228. [[CrossRef](#)]
19. Sun, J.; Di, L.; Sun, Z.; Shen, Y.; Lai, Z. County-Level Soybean Yield Prediction Using Deep CNN-LSTM Model. *Sensors* **2019**, *19*, 4363. [[CrossRef](#)]
20. Yu, N.; Li, L.; Schmitz, N.; Tian, L.F.; Greenberg, J.A.; Diers, B.W. Development of methods to improve soybean yield estimation and predict plant maturity with an unmanned aerial vehicle based platform. *Remote Sens. Environ.* **2016**, *187*, 91–101. [[CrossRef](#)]
21. Taha, R.S.; Seleiman, M.F.; Alotaibi, M.; Alhammad, B.A.; Rady, M.M.; Mahdi, A.H.A. Exogenous potassium treatments elevate salt tolerance and performances of *Glycine max* L. by boosting antioxidant defense system under actual saline field conditions. *Agronomy* **2020**, *10*, 1741. [[CrossRef](#)]

22. Conab—Monitoramento Agrícola. Available online: <https://www.conab.gov.br/index.php/info-agro/safras/graos/monitoramento-agricola> (accessed on 15 April 2021).
23. Wei, M.; Molin, J. Soybean Yield Estimation and Its Components: A Linear Regression Approach. *Agriculture* **2020**, *10*, 348. [[CrossRef](#)]
24. Khanal, S.; Fulton, J.; Klopfenstein, A.; Douridas, N.; Shearer, S. Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. *Comput. Electron. Agric.* **2018**, *153*, 213–225. [[CrossRef](#)]
25. Soltis, P.S.; Nelson, G.; Zare, A.; Meineke, E.K. Plants meet machines: Prospects in machine learning for plant biology. *Appl. Plant Sci.* **2020**, *8*, e11371. [[CrossRef](#)]
26. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [[CrossRef](#)]
27. Ball, J.E.; Anderson, D.T.; Chan, C.S. Comprehensive survey of deep learning in remote sensing: Theories, tools, and challenges for the community. *J. Appl. Remote Sens.* **2017**, *11*, 042609. [[CrossRef](#)]
28. Castro, W.; Junior, J.M.; Polidoro, C.; Osco, L.P.; Gonçalves, W.; Rodrigues, L.; Santos, M.; Jank, L.; Barrios, S.; Valle, C.; et al. Deep Learning Applied to Phenotyping of Biomass in Forages with UAV-Based RGB Imagery. *Sensors* **2020**, *20*, 4802. [[CrossRef](#)]
29. Chen, Y.; Ribera, J.; Boomsma, C.; Delp, E.J. *Plant Leaf Segmentation for Estimating Phenotypic Traits*; Video and Image Processing Laboratory (VIPER); Purdue University: West Lafayette, IN, USA, 2017; pp. 3884–3888.
30. Jin, Z.; Azzari, G.; You, C.; Di Tommaso, S.; Aston, S.; Burke, M.; Lobell, D.B. Smallholder maize area and yield mapping at national scales with Google Earth Engine. *Remote Sens. Environ.* **2019**, *228*, 115–128. [[CrossRef](#)]
31. Nevavuori, P.; Narra, N.; Lipping, T. Crop yield prediction with deep convolutional neural networks. *Comput. Electron. Agric.* **2019**, *163*, 104859. [[CrossRef](#)]
32. Barbosa, A.; Trevisan, R.; Hovakimyan, N.; Martin, N.F. Modeling yield response to crop management using convolutional neural networks. *Comput. Electron. Agric.* **2020**, *170*, 105197. [[CrossRef](#)]
33. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)]
34. Finoto, E.L.; Sediyaama, T.; Carrega, W.C.; De Albuquerque, J.A.A.; Cecon, P.R.; Reis, M.S. Efeito da aplicação de fungicida sobre caracteres agrônômicos e severidade das doenças de final de ciclo na cultura da soja. *Rev. Agro@Mambiente On-Line* **2011**, *5*, 44. [[CrossRef](#)]
35. Masuka, B.; Atlin, G.N.; Olsen, M.; Magorokosho, C.; Labuschagne, M.; Crossa, J.; Bänziger, M.; Pixley, K.V.; Vivek, B.S.; von Biljon, A.; et al. Gains in Maize Genetic Improvement in Eastern and Southern Africa: I. CIMMYT Hybrid Breeding Pipeline. *Crop. Sci.* **2017**, *57*, 168–179. [[CrossRef](#)]
36. Morrison, M.J.; Voldeng, H.D.; Cober, E.R. Agronomic Changes from 58 Years of Genetic Improvement of Short-Season Soybean Cultivars in Canada. *Agron. J.* **2000**, *92*, 780–784. [[CrossRef](#)]
37. Jin, J.; Liu, X.; Wang, G.; Mi, L.; Shen, Z.; Chen, X.; Herbert, S.J. Agronomic and physiological contributions to the yield improvement of soybean cultivars released from 1950 to 2006 in Northeast China. *Field Crop. Res.* **2010**, *115*, 116–123. [[CrossRef](#)]
38. Todeschini, M.H.; Milioli, A.S.; Rosa, A.C.; Dallacorte, L.V.; Panho, M.C.; Marchese, J.A.; Benin, G. Soybean genetic progress in South Brazil: Physiological, phenological and agronomic traits. *Euphytica* **2019**, *215*, 124. [[CrossRef](#)]
39. Xie, C.; Yang, C. A review on plant high-throughput phenotyping traits using UAV-based sensors. *Comput. Electron. Agric.* **2020**, *178*, 105731. [[CrossRef](#)]
40. Eugenio, F.C.; Grohs, M.; Venancio, L.P.; Schuh, M.; Bottega, E.L.; Ruoso, R.; Schons, C.; Mallmann, C.L.; Badin, T.L.; Fernandes, P. Estimation of soybean yield from machine learning techniques and multispectral RPAS imagery. *Remote Sens. Appl. Soc. Environ.* **2020**, *20*, 100397. [[CrossRef](#)]