BIOSCIENCE JOURNAL

BIOPHYSICAL CHARACTERISTICS OF SOYBEAN ESTIMATED BY REMOTE SENSING ASSOCIATED WITH ARTIFICIAL INTELLIGENCE

Franciele Morlin CARNEIRO[1], Mailson Freire de OLIVEIRA[2], Samira Luns Hatum de ALMEIDA[3], Armando Lopes de BRITO FILHO[4], Carlos Eduardo Angeli FURLANI[5], Glauco de Souza ROLIM[5], Antonio Sergio FERRAUDO[5], Rouverson Pereira da SILVA[5]

[1] School of Plant, Environmental and Soil Sciences, Louisiana State University, Baton Rouge, Louisiana, United States of America.
[2] Crop, Soil & Environmental Sciences Department, Auburn University, Auburn, Alabama, United States of America.
[3] Postgraduate program in Agronomy (Crop Production), School of Agricultural and Veterinarian Sciences, São Paulo State University, Jaboticabal, São Paulo, Brazil.
[4] Postgraduate program in Agronomy (Soil Science), School of Agricultural and Veterinarian Sciences, São Paulo State University, Jaboticabal, São Paulo, Brazil.
[5] Engineering and Exact Sciences Department, School of Agricultural and Veterinarian Sciences, São Paulo State University, Jaboticabal, São Paulo, Brazil.

Corresponding author:
Name: Franciele Morlin Carneiro
Email: franmorlin1@gmail.com

How to cite: CARNEIRO, F.M., et al. Biophysical characteristics of soybean estimated by remote sensing associated with artificial intelligence. *Bioscience Journal*. 2022, **38**, e38024. https://doi.org/10.14393/BJ-v38n0a2022-55925

Abstract
The biophysical characteristics of vegetative canopies, such as biomass, height, and canopy diameter, are of paramount importance for the study of the development and productive behavior of crops. Faced with a scarcity of studies aimed at estimating these parameters, the objective of this study was to evaluate the performance of artificial neural networks (ANNs) applied to Proximal Remote Sensing (PRS) to estimate biophysical characteristics of soybean culture. The data used to train and validate the ANNs came from an experiment composed of 65 plots with 30 x 30 m mesh, its development was carried out in the 2016/2017 crop in the Brazilian agricultural area. The evaluations were carried out at 30, 45, 60, and 75 days after sowing (DAS), monitoring the spatial and temporal variability of the biophysical characteristics of the soybean crop. Vegetation indexes were collected using canopy sensors. The accuracy and precision were determined by the coefficient of determination (R2) and the error of the forecasts by MAPE (Mean Absolute Percentage Error). PRS and ANNs showed high potential for application in agriculture, since they obtained good performance in the estimation of height (R2 = 0.89) and canopy diameter (R2 = 0.96), being fresh biomass (R2 =0.98) and dry biomass (R2 = 0.97) were the best-estimated variables.

Keywords: Artificial Neural Networks. Active Optical Sensor. *Glycine max* L. Machine Learning. Vegetation Index.

1. Introduction

The search for more sustainable and cost-saving agriculture has contributed to the advancement of technologies in the area and the improvement of new methodologies. Nowadays, information technology has become very important in several systems, making it possible to manage the harvest, optimize the application of the products, and manage the planting. In this context, bioinformatics techniques such as Artificial Neural Networks (ANNs) prove to be relevant tools in several areas of agriculture with great application in the analysis of agricultural variables.

Artificial neural networks (ANNs) are computational models characterized by artificial neurons connected by a large number of interconnections called artificial synapses (Silva et al. 2016). They are used in the most diverse areas with potential for application in forecasting systems, whose objective is to estimate future values of a process taking into account several previous measures observed. Among the available applications are climate forecasts, time series forecasts, financial market forecasts, among others (Silva et al. 2016).

One of the most relevant characteristics of ANNs is their ability to learn from examples. ANNs have shown greater performance due to factors such as efficiency in learning and generalization, making them capable of solving complex problems; enable modeling with numerical and categorical variables; parallel distributed structure (layers) and robust; they are tolerant of outliers or outliers, they can model different variables and non-linear relationships (Haykin 2001).

Using crop simulation models, ANNs can be used with great success in various situations for crop estimates and forecasts (Andrade Júnior et al. 2006). Crop growth models exhibit potential use for crop planning and management, helping to understand physiological, environmental, and genetic interactions, as well as decisions on cultural practices, such as irrigation and fertilizer distribution (Boote et al. 1996).

Another concept that has contributed to significant advances in the agricultural environment is remote sensing (RS). Remote sensing is a science-based on the interaction of electromagnetic radiation with the Earth's surface and allows obtaining data from a target without direct contact between the sensor and the target. As the main advantage, mainly concerning the plant study, the SR allows the extraction of data from the vegetation without the need for destructive samples. The levels of data collection in this science can be orbital, aerial, and terrestrial, the terrestrial being the one that allows the study of the target more punctually, contributing to a better understanding of the behavior of this target in the face of electromagnetic radiation.

There is still a lack of work in the literature with the use of artificial intelligence and PRS (Proximal Remote Sensing), highlighting the relevance of this research for understanding the temporal and spatial analysis related to the biophysical characteristics of the plant, such as plant biomass, height, canopy width, and productivity. The estimation of biophysical characteristics is of utmost importance since it can contribute to the reduction of labor, monitoring of the vigor of cultivation, and destination of the final product on the market.

For the prediction of crop production, the formulation of a mathematical model is limited and difficult, due to the non-linearity of the data of the parameters related to it and to the complexity (Braga et al. 2012). Thus, several authors recommend the use of RNAs when systems are complex (Vieira et al. 2009; Jana and Mohanty 2012). Given the exhibited, the objective of this work was to evaluate the performance of the artificial neural network applied to remote sensing at the terrestrial level with canopy sensors to estimate biophysical characteristics of soybean, an important crop in the world commodities market.

2. Material and Methods

Description of the experimental area

The experiment was carried out in an agricultural area of the municipality of Jaboticabal located at coordinates 21º15'19,6''S and 48º15'38,5''W, state of São Paulo, Brazil (Figure 1). The area has as characteristic soil the RED LATOSOL (Santos et al. 2018). According to the Köppen climatic classification, the region has a climate characterized as AW, dry winter tropical with an average temperature of 22.2ºC (Alvares et al. 2013).

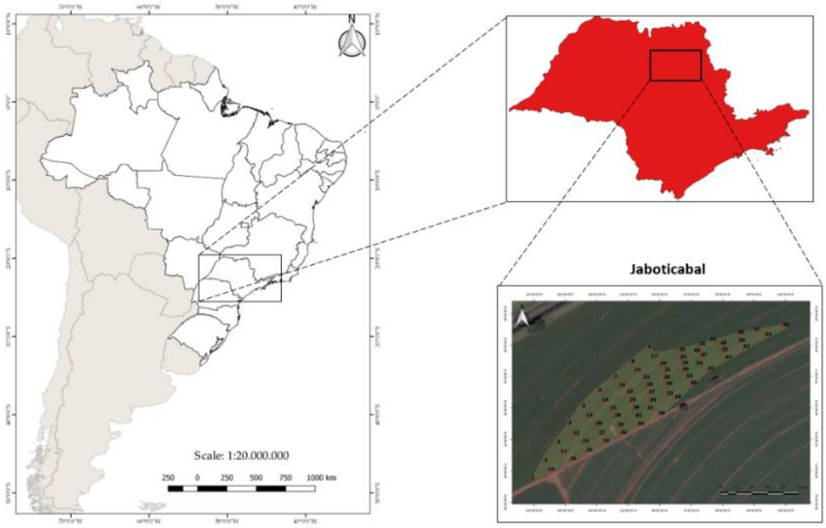

Figure 1. The experimental area is located in Jaboticabal, state of São Paulo, Brazil.

Data acquisition

The data used to train and validate the ANNs came from a soybean experiment conducted in the 2016/2017 harvest, consisting of 65 sample points with 30 x 30 m mesh. The evaluations were performed at 30, 45, 60, and 75 days after sowing (DAS), stages V3, V5, V6, and R4 respectively, in which fresh and dry biomass, canopy width, chlorophyll index, and plant height were evaluated.

Each sample point was consisted of two lines 5 meters long with 0.45 m of spacing between lines, making 4.5 m² of usable area per point. All sampling points were georeferenced with the GNSS receiver - Global Navigation Satellite System - Trimble R6, with the GNSS system, receiving RTK positioning signal - Real-Time Kinematic - and high precision antenna (Trimble 2013; Embratop 2017). In addition to the variables mentioned above, vegetation index data (NDVI - Normalized Difference Vegetation Index; NDRE - Normalized Difference Red Edge; and IRVI - Inverse Ratio Vegetation Index) were collected using canopy sensors. The readings of the vegetation indexes were performed on the canopy of the plant according to Grohs et al. (2011).

Vegetation index

The sensors used to obtain the vegetation indexes were GreenSeeker and OprTX. The GreenSeeker model 500, Trimble, emitted radiation at the near-infrared wavelengths - NIR (770 nm) and red (660 nm), with a bandwidth of about 25 nm (Povh et al. 2008; Amaral et al. 2015). The reflected light is captured by the sensor to calculate the NDVI vegetation index (Motomiya et al. 2014) and the IRVI. The readings with the sensor were performed at a working height of 0.6 to 0.7 m.

The OptRX active sensor ACS430 model, Ag Leader, allows the obtainment of NDRE and NDVI vegetation indices. However, in this work, the NDVI collected by the GreenSeeker sensor was considered. Table 1 shows the calculation for the evaluated vegetation indices.

Table 1. Vegetation indexes.

Vegetation index	Index calculation	Source
NDVI	$\text{NDVI} = \frac{F_{NIR}-F_{Red}}{F_{NIR}+F_{Red}}$	Rouse et al, (1973)
NDRE	$\text{NDRE} = \frac{\text{NIR-RE}}{\text{NIR+RE}}$	Buschmann and Nagel (1993)
IRVI	$\text{IRVI} = \frac{R_{650}}{R_{770}}$	Kapp Júnior et al, (2016)

NDVI Normalized Differential Vegetation Index, *NIR* emission fractions in near-infrared, Red emission fractions of red, *NDRE* Normalized Difference Red Edge, *RE* red-edge rate indices, *IRVI* Inverse Ratio Vegetation Index, *NIR* 774 nm, *Red* 656 nm, *RE* 720 nm.

Quality indicators

The evaluated parameters were plant height and width, fresh biomass, and chlorophyll index. For the height of the plant and the canopy width, a measuring tape graduated in centimeters was used. To obtain fresh biomass ('in natura"), an area of 0.25 m2 (0.5 x 0.5 m) was used (Figure 2). The plants inside this frame were cut close to the ground and then weighed on a semi-analytical scale to measure the fresh weight. The samples were placed in an oven with a circulation at 65ºC for 72 h (Gobbi et al. 2009; Grohs et al. 2009) to obtain the dry mass.

Figure 2. Collection of fresh biomasses using the frame with dimensions 0.5 x 0.5 m.

The Chlorophyll Index was acquired by the Marcone® model CCM-200 plus chlorophyll meter, with an accuracy of ± 1 CCI unit (Chlorophyll Content Index). Readings were performed by randomly collecting three leaves per plot and three readings per leaf.

Artificial neural networks: models design

The multilayer ANNs, artificial neuron systems, have a basic structure of an input layer, two hidden layers, and an output layer, as illustrated in Figure 3.

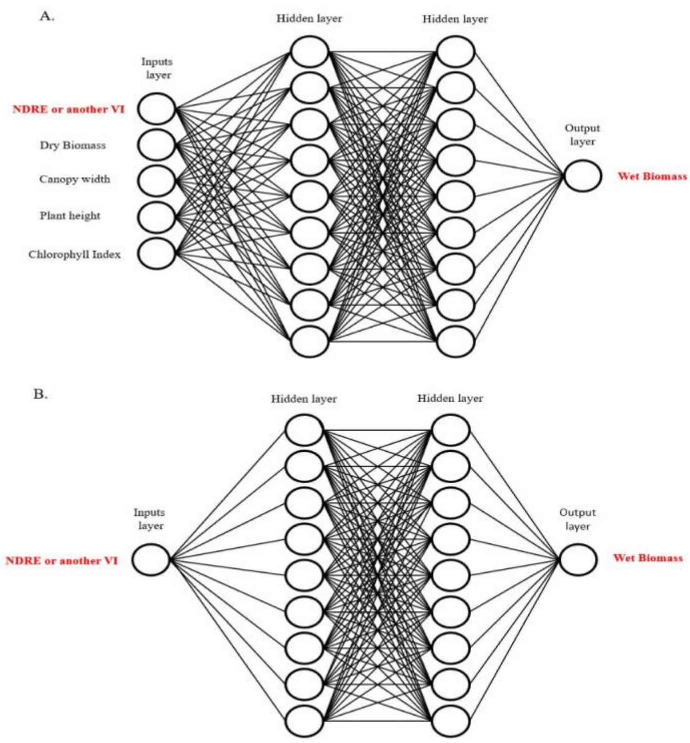

Figure 3. Multilayer feedforward artificial neural network architecture with topologies: (A) 5 → 8 → 8 → 1; and (B) 1 → 8 → 8 → 1.

In the input layer, the information used to make the prediction of the output layer was variable. In the hidden layers, eight neurons were used and for the output layer only one neuron. Multilayer networks were trained using the supervised learning algorithm for the backpropagation of errors (Backpropagation). The backpropagation neural network is trained with the inputs adjusted to the output variables in two phases (Zhang 2015). The networks were interconnected by connecting forces represented by values that are called synaptic weights, which are responsible for storing acquired knowledge. The values used in the input layers were normalized according to Equation 1.

$$y_i = \frac{x_i - x_{max}}{x_{max} + x_{min}} \qquad (1)$$

where, in (1):

Yi = input vector value (example: mean chlorophyll)

xmin = minimum value

xmax = maximum value

The output value of each neuron in bed k is expressed as yk = g (ak).

where:

g = ak activation function

ak = synaptic function, which is a linear combination of normalized input values and synaptic weights as shown in Equation 2.

$$a_k = \sum_j y_j w_{kj} \qquad (2)$$

where in (2):

wkj = synaptic weights linking the yj input values with each k neuron.

The transfer or activation function in the neurons of each hidden layer was the hyperbolic function, shown in Equation 3.

$$g(a_k) = \frac{e^{ak} -- e^{-ak}}{e^{ak} + e^{-ak}} \qquad (3)$$

where in (3):

e = Neperian algorithm

Training and validation models

For the training and validation of the models, the database was divided into 80% of the data for training and 20% for validation (Silva et al. 2016). Because the information was collected at different phenological stages, this percentage of the database partition was applied to each stage of data collection.

The training and validation procedures for neural models were implemented in the Neural Networks package of the Statistica data analysis software (Statistica 7.0, Statsoft Inc, Tulsa, OK). In this work, eight neurons, and four layers were used as architecture, comprising: an initial layer, two intermediates, and an output layer, which can be seen in Figure 3.

Linear was used in the input and output layer and the following variables were used: fresh and dry biomass (kg ha^{-1}), canopy width (cm), plant height (cm), and chlorophyll index (CCI). For architecture 5 → 8 → 8 → 1 (Figure 3A), all morphological variables plus the vegetation index of interest were used in the input layer, that is, an index plus all variables, however, when the architecture was 1 → 8 → 8 → 1 (Figure 3B) only one vegetation index was used.

In the development of ANN in the intermediate or hidden layer, the function of activation was the sigmoid Hyperbolic Tangent and the network was trained 1000 times because the free parameters are obtained at random Soares et al. 2014.

Performance test

In the development of ANN in the intermediate or hidden layer, the function of activation was the sigmoid Hyperbolic Tangent and the network was trained 1000 times because the free parameters are obtained at random (Soares et al. 2014).

$$MAPE=\frac{\sum_{i=1}^{n}\left(\left|\frac{Yest_i-Yobs_i}{Yobs_i}\right|*100\right)}{n} \quad (4)$$

where in (4):
n = data numbers
Yesti = value of the variable estimated by the network
Yobsi = observed variable value

$$R^2=\frac{SQR}{SQT} \quad (5)$$

Where in (5):
SQR = sum of the regression squares
SQT = sum of total squares

3. Results

Artificial neural networks (ANNs)

The results of this study are divided into two main sections: (i) artificial neural networks and (ii) performance of the network architecture with the estimated and observed data. Table 2 shows the results of the accuracy and precision of the neural model validation step to estimate the variables.

Table 2. Accuracy (MAPE) and precision (R^2) of the validation of neural models to estimate the variables fresh and dry biomass, canopy width, plant height, and chlorophyll index for soybean culture.

	Fresh biomass		Dry biomass		Canopy width		Height of plant		Chlorophyll index	
	R^2	MAPE (%)	R^2	MAPE (%)	R^2	MAPE (%)	R^2	MAPE (%)	R^2	MAPE (%)
TV_NDRE	0.97	14	0.97	13	0.94	10	0.79	10	0.30	17.16
NDRE	0.88	23	0.86	21	0.94	14	0.89	11	0.29	17.40
TV_NDVI	0.98	12	0.97	10	0.96	8	0.79	11	0.12	22.37
NDVI	0.84	21	0.70	27	0.84	15	0.73	13	0.04	20.75
TV_IRVI	0.97	13	0.95	13	0.95	9	0.59	13	0.10	23.54
IRVI	0.82	32	0.78	30	0.83	15	0.71	14	0.03	22.33

TV: all variables; NDRE: Normalized Difference Red Edge Index; NDVI: Normalized Difference Vegetation Index; IRVI: Reverse Vegetation Index.

ANNs performance

The efficiency of the networks was analyzed in performance graphs, which can be seen in Figure 4, between the relationship between the NDRE data and the dry biomass.

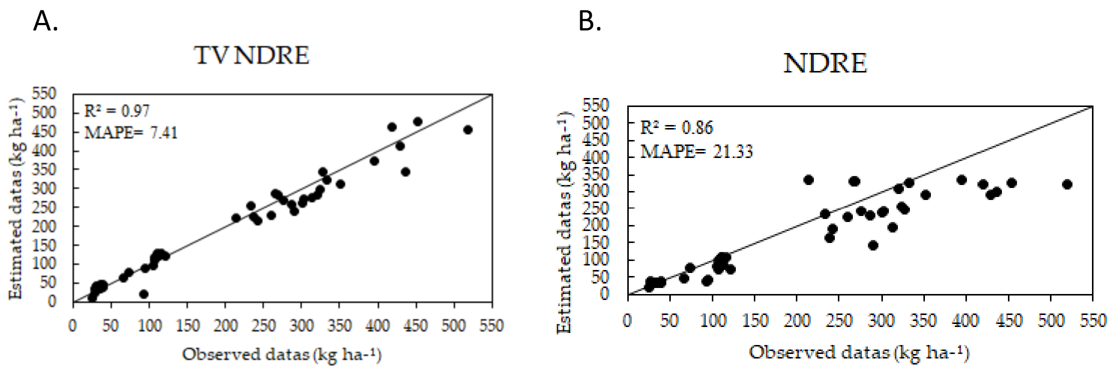

Figure 4. Scatterplots indicating 1:1 relationship with NDRE and dry biomass from the best-trained neural network. A – all variables plus NDRE; B – NDRE.

In Figure 5, comparative analyzes were performed between the predicted data and the real ones, verifying that the behavior of the prediction curve behaves similarly with the real data collected in the field, mainly, in Figure 5A it presented this behavior. As was said and observed in Figure 4A and Table 2, the highest R^2 was when they had the largest amounts of input layers.

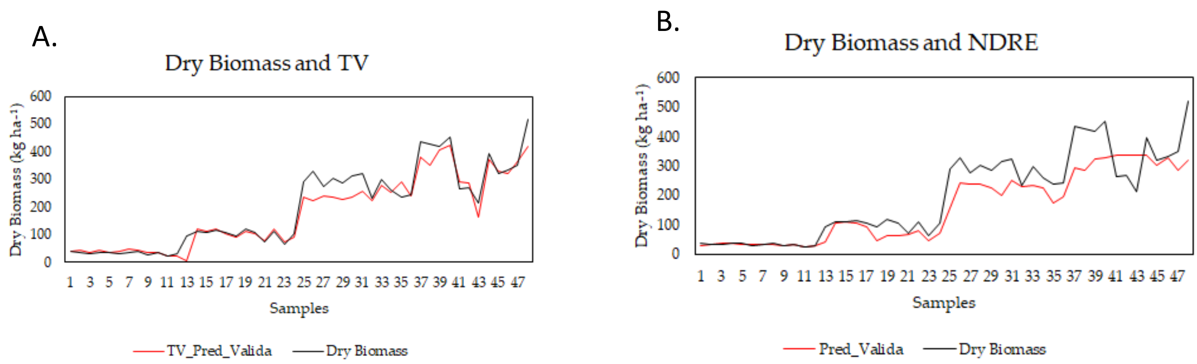

Figure 5. Comparative analyzes between: A – predicted dry biomass with all variables (TV) versus B – actual dry biomass with NDRE.

In Figures 6A and 7A, it was verified that the architecture 5 → 8 → 8 → 1 demonstrated better predictive capacity for the variable canopy width due to the accuracy and precision results of the validation of neural models to estimate the variables by means of R^2 and MAPE. Since in Figure 7A, the predicted data for this same architecture proved to be very close to the real data.

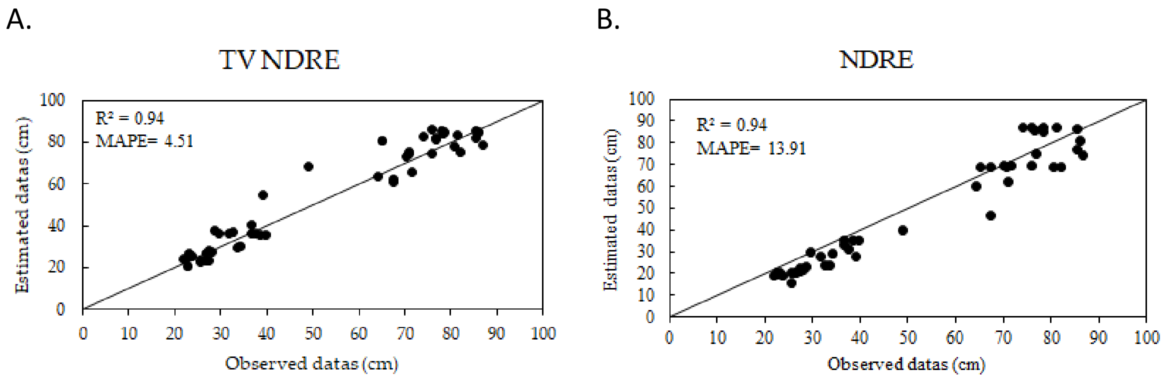

Figure 6. Scatterplots indicating 1:1 relationship with NDRE and canopy width from the best-trained neural network. A – all variables plus NDRE; B – NDRE.

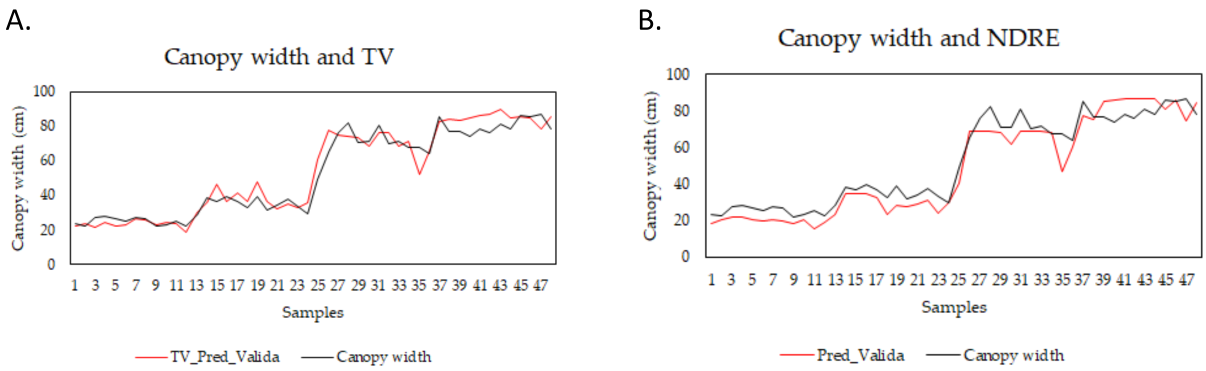

Figure 7. Comparative analysis between: A – predicted canopy widths with all variables (TV) versus B – actual canopy width with NDRE.

In Figures 8 and 9, the results obtained for the plant height variable differed from the others because the architecture that showed the best performance was 1 → 8 → 8 → 1 with a determination coefficient of 0.89 (Figure 8B), however, the MAPE between the two architectures evaluated for this variable differed only by 0.39. For this reason, which in Figure 9A pointed out a greater similarity between the predicted data and the actual data due to the accuracy and precision of the validation.

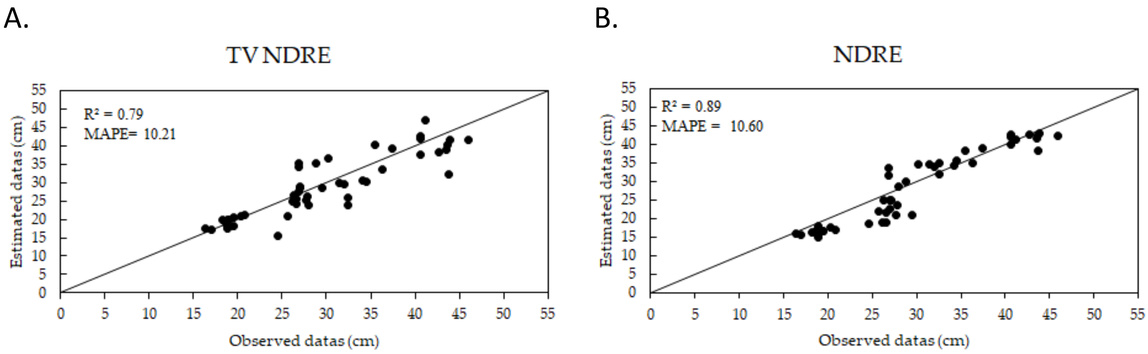

Figure 8. Scatterplots indicating 1:1 relationship with NDRE and height from the best-trained neural network. A – all variables plus NDRE; B – NDRE.

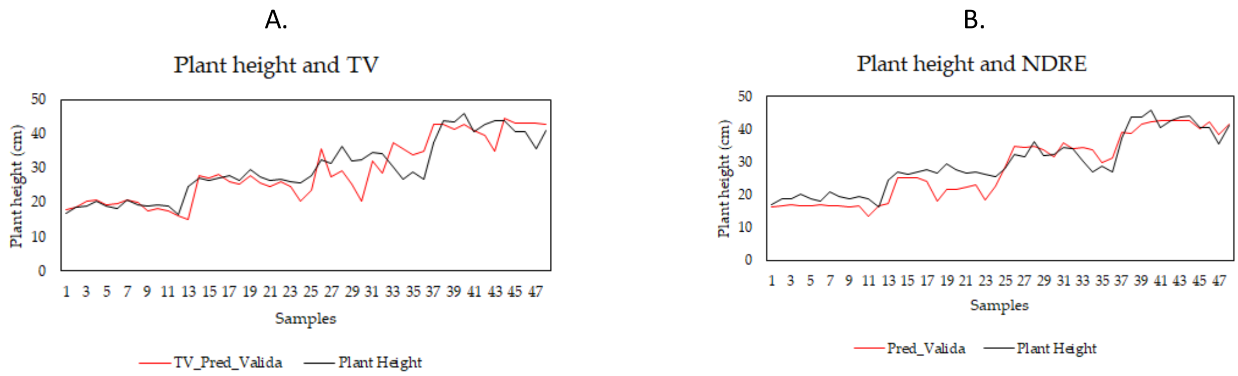

Figure 9. Comparative analysis between: A – predicted plant height with all variables (TV) versus B – actual plant height with NDRE.

4. Discussion

Artificial neural networks (ANNs)

It was demonstrated that most of the data obtained the determination coefficients (R^2) greater than 0.70 (Table 2), which according to Garcia (1989) values greater than 0.70 point to good adjustments. The forecast errors were calculated using the MAPE error measure (Mean Absolute Percentage Error) and it was possible to observe that the smallest errors were for architecture 5 → 8 → 8 → 1 (Figure 3A), demonstrating greater accuracy when compared to architecture 1 → 8 → 8 → 1 (Figure 3B).

As seen in Table 2, when all morphological variables were used, the results of both precision and accuracy demonstrated the best performance for the architecture 5 → 8 → 8 → 1. For the field collection of these variables in large areas it requires a lot of hands - when the operation is carried out manually, being very laborious, therefore, most of the work carried out with remote sensing in large areas is done with the use of ARP (Remotely Piloted Aircraft) or satellite with passive sensors, however, is made with the use of tractors facilitates and optimizes operation at the level of land collection. However, data collected at the terrestrial level as performed in this work have less interference from atmospheric conditions because the sensor is closer to the target object and because it produces its own source of electromagnetic radiation, with reductions in errors during collection. Novo (2008) cites those that atmospheric factors affect in the collection of data reflectances, such as water vapor, clouds, carbon dioxide, ozone, and oxygen.

Passive sensors need natural radiation to obtain the data. However, the active sensors produce their own radiation, for example, a photographic camera for producing its flash is considered an active sensor, however, the camera that needs sunlight is a passive sensor (Navalgund 2002).

Remote sensing has shown great potential in the area of Precision Agriculture as it is a non-destructive method, obtaining data in real-time, helps the farmer in monitoring the growth stages of the crop through the use of sensors, which, which generate vegetation indexes that assist in monitoring the vigor of the plant. These advantages were also observed by Jensen (1996), Casanova et al. (1998), Alt et al. (2000), Diker and Bausch (2003), Hansen and Schjoerring (2003) and others.

As seen in Table 2, most of the models have high predictive capacity because they had an R^2 greater than 0.90, and this was also observed by Soares et al. 2015, to estimate corn production using artificial neural networks, in which they highlighted the strong correlation (R^2 of 1.00) for the grain production variable.

Through the determination coefficient, the explanatory power of the estimates generated by the models (Y) and the values measured in the field (X) are verified, that is, it evaluates the quality of the adjustment between the variables when this coefficient is equal to zero, it means that the explained variation of Y is zero with the adjusted line parallel to the X-axis. However, if the R^2 is equal to one, the line will explain the entire variation of Y. The higher the values of this coefficient better the quality (Toledo and Ovalle 1995).

Since the highest values of R^2 were for biomass showing how much the vegetation indexes have a high relationship with this variable, demonstrating that biomass was the best variable predicted through remote sensing. Corroborating the obtained result, Grohs et al. (2009) found that when there was an increase in biomass, the NDVI value also increased until the saturation of this index.

For the variables canopy width and plant height, they also showed high coefficients indicating high relations with the vegetation indexes, mainly for architecture 5 → 8 → 8 → 1. However, for the variable plant height, the coefficients were higher only for an input variable, that is, for architecture 1 → 8 → 8 → 1. As for the chlorophyll index variable, it did not present high values of the coefficients of determination, demonstrating low prediction through neural networks using only this variable.

And these different results obtained can be explained due to the factors that interfere in the spectral responses of the canopies of the plants, as observed by Formaggio and Sanches (2017), which can be mainly due to the structure of the canopy (plant architecture and planting density), due to the geometry aspects of scene and lighting, and soils (substrates). The same authors mention that in the initial growth stages of the crop there is the spectral domain of the soil, however, during vegetative development/flowering there is the spectral domain of the green cover, and finally, in maturation and senescence, the spectral domain occurs soil and dry vegetation.

The NDVI differed from the IRVI and the NDRE, probably due to the saturation problem that this index had, due to its greater sensitivity, which was also verified by Cao et al. (2015) and Lu et al. (2017).

ANNs performance

It was verified when all the morphological variables are used, for architecture 5 → 8 → 8 → 1 (Figure 4A), the estimates were closer to the 1: 1 line due to the higher coefficient of determination (R^2) (0.97) and MAPE was lower (7.41) when compared to just one entry, for architecture 1 → 8 → 8 → 1 (Figure 4B), in which case the entry was NDRE, and that too it was presented in Table 2 in which the highest R^2 were when all variables were used as input. In addition, for architecture 1 → 8 → 8 → 1 it demonstrated how close the estimate was to the estimated data (Figure 4A).

An accuracy (MAPE) greater than 10% has been achieved (Table 2), this means, for example, if the model is estimating a plant height of 100 cm, this value in the field (observed value) can be 110 or 90 cm. To increase the accuracy of the models (low down MAPE), perhaps adding other variables to the model that can explain other components of the studied phenomena. Other variables predicted by remote sensings, such as soybean productivity, can be better predicted when using other components in the models, such as elevation and flow accumulation (Kross et al. 2020). Therefore, this approach can increase the accuracy of the predictions of the variables we studied. It was also observed that, when we use only the vegetation index, less accuracy was found when compared with models that use vegetation indexes and covariates as input, this result is logically explained by the fact that when we have other variables that explain the predicted phenomena the model has more information related to the predicted variable, which can increase the accuracy of the forecast. In our study, the covariables used were manually measured variables, which means that it is unreasonable to use models that require manual variables, it is preferable to use models that use only remote sensing data (NDRE). These models showed reasonable accuracy and can be used to generate estimated variability maps within the field of canopy width, dry biomass, canopy width, and plant height.

Barros (2005) found that a trained network with a small number of connections is unable to make the most of its potential, however, when there is an excess of connections it can adapt to noise and affect its training. Thus, it was verified in the present work for the architecture 5 → 8 → 8 → 1 demonstrated that the greater the number of layers entered, the values of R^2 are higher when compared with only one input layer, for architecture 1 → 8 → 8 → 1.

From 60 DAS, it was observed in sample 25 that the model underestimated the estimates, which may be related to the increase in the development of biomass affecting the reading of the vegetation index, demonstrating that up to 45 DAS the model had better behavior when compared to the estimated real data (Figure 5).

Soares et al. (2014), proposed a methodology for estimating the water retention curve in the soil by means of artificial neural networks, saw that with the increase in the number of neurons in the hidden layer, there was no decrease in the average relative error, however, this reduction happened in input layer with the increasing number of variables. This was also shown in this work, and it can be seen in Figure 8A that the behavior of the curves was similar, demonstrating the high predictive power of the dry biomass variable, and besides, how much this variable correlates with the vegetation indexes, especially with the NDRE.

As verified by Soares et al. (2015) for the estimation of corn grain productivity, observed that the results obtained through the statistical analysis of artificial neural networks are feasible and can be used as a modeling tool to estimate productivity, making it a relevant tool for the agricultural sector. Mutanga and Skidmore (2004) saw in their research that neural networks obtained better results than when it was used, traditional statistics, linear regression for the prediction of savannah grass quality.

It was found when morphological information was used in the input layer, the predictions were the best in both accuracy and precision. However, when using information without physical contact with the plant, only the vegetation index (NDRE), more parsimonious models were obtained to estimate the variables plant height, canopy width, and soybean biomass.

As noted, the active sensors can be used in the prediction of variables with the possibility of using the estimates generated by the model for recommendations and investigations based on the spatial and temporal variability of the morphological parameters and can also be used to monitor the vigor of the culture within the field, allowing greater knowledge of the crop.

Estimates made from trained and validated neural models can support the direction of investigations, whether manual or robotic, and can also contribute to the application of phytosanitary products or fertilizers using aircraft or machines for localized applications, contributing to the management in a specific site.

New studies with ANNs should be done, mainly, aimed at the identification of new patterns such as, different agricultural cultures, other vegetation indexes in addition to sensors that have different functioning mechanisms and radiometric, spatial, and spectral resolution as sensors embedded in aircraft remotely piloted.

5. Conclusions

The ANNs demonstrated good performance to predict the morphological characteristics of the crop, such as biomass, plant height, and canopy width. Dry biomass stood out as the best variable predicted through the SRP and the networks. With the adoption of remote sensing and ANNs in agriculture, they proved to be an excellent tool for predicting variables. The artificial neural networks were able to identify the patterns between the information from the proximal remote sensors and the morphological variables, this method being indicated to predict the height, width, fresh and dry mass of the soybean crop.

Authors' Contributions: CARNEIRO, F.M.: conception and design, acquisition of data, analysis and interpretation of data, drafting the article, and critical review of important intellectual content; OLIVEIRA, M.F.: conception and design, acquisition of data, analysis and interpretation of data, drafting the article, and critical review of important intellectual content; ALMEIDA, S.L.H.: conception and design, acquisition of data, analysis and interpretation of data, drafting the article, and critical review of important intellectual content; BRITO FILHO, A.L.: conception and design, acquisition of data, analysis and interpretation of data, drafting the article, and critical review of important intellectual content; FURLANI, C.E.A.: conception and design, acquisition of data, analysis and interpretation of data, drafting the article, and critical review of important intellectual content; ROLIM, G.S.: conception and design, acquisition of data, analysis and interpretation of data, drafting the article, and critical review of important intellectual content; FERRAUDO, A.S.: conception and design, acquisition of data, analysis and interpretation of data, drafting the article, and critical review of important intellectual content; SILVA, R.P.: conception and design, acquisition of data, analysis and interpretation of data, drafting the article, and critical review of important intellectual content. All authors have read and approved the final version of the manuscript.

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

Ethics Approval: Not applicable.

Acknowledgments: The authors would like to thank the funding for the realization of this study provided by the Brazilian agency CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brasil), Finance Code 142367/2015-0.

References

ALT, C., KAGE, H. and STÜTZEL, H. Modelling nitrogen content and distribution in cauliflower (Brassica oleracea L. botrytis). *Annals of Botany.* 2000, **86**(5), 963-973. https://doi.org/10.1006/anbo.2000.1252

ALVARES, C.A., et al. Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*. 2013, **22**(6), 711-728. https://doi.org/10.1127/0941-2948/2013/0507

AMARAL, L.R., et al. Comparison of crop canopy reflectance sensors used to identify sugarcane biomass and nitrogen status. *Precision Agriculture.* 2015, **16(**1), 15-28. https://doi.org/10.1007/s11119-014-9377-2

ANDRADE JÚNIOR, A.S., FIGUEREDO JÚNIOR., L.G.M., CARDOSO, M.J. and RIBEIRO, V.Q. Parametrização de modelos agrometeorológicos para estimativa de produtividade da cultura do milho na região de Parnaíba, Piauí. *Revista Ciência Agronômica*. 2006, **37**(2), 130–134.

BARROS, A.C.A. *Otimização de redes neurais para previsão de séries temporais*. Recife: Escola Politécnica de Pernambuco – Universidade de Pernambuco, Trabalho de Conclusão de Curso, 2005.

BOOTE, K.J., JONES, J.W. and PICKERING, N.B. Potential uses and limitations of crop models. *Agronomy Journal*. 1996, **88**(5), 704-716. https://doi.org/10.2134/agronj1996.0002196200-8800050005x

BRAGA, A.P., CARVALHO, A.P.L.F. and LUDERMIR, T.B. *Redes Neurais Artificiais:* teoria e aplicações. 2nd ed. Rio de Janeiro: LTC, 2012.

CAO, Q., et al. Active canopy sensing of winter wheat nitrogen status: An evaluation of two sensor systems. *Computers and Electronics in Agriculture*. 2015, **112**, 54-67. https://doi.org/10.1016/j.compag.2014.08.012

CASANOVA, D., EPEMA, G.F. and GOUDRIAAN, J. Monitoring rice reflectance at field level for estimating biomass and LAI. *Field Crops Research*. 1998, **55**(1-2), 83-92. https://doi.org/10.1016/S0378-4290(97)00064-6

DIKER, K. and BAUSCH, W.C. Potential use of nitrogen reflectance index to estimate plant parameters and yield of maize. *Biosystems Engineering*. 2003, **85**(4), 437-447. https://doi.org/10.1016/S1537-5110(03)00097-7

EMBRATOP. *Receptor GPS Trimble R6 (L1/L2)*. 2017. Available from: http://www.embratop.com.br/produto/receptor-gps-trimble-r6-l1-l2/

FORMAGGIO, A.R. and SANCHES, I.D. *Sensoriamento Remoto em agricultura.* São Paulo: Oficina de Textos, 2017.

GARCIA, C.H. *Tabelas para classificação do coeficiente de variação*. Piracicaba: Instituto de Pesquisas e Estudos Florestais, Escola Superior de Agricultura "Luiz de Queiroz", 1989.

GOBBI, K.F., et al. Características morfológicas, estruturais e produtividade do capim-braquiária e do amendoim forrageiro submetidos ao sombreamento. *Revista Brasileira de Zootecnia*. 2009. **38**(9), 1645-1654. https://doi.o-rg/10.1590/S1516-35982009000900002

GROHS, D.S., BREDEMEIER, C., MUNDSTOCK, C.M. and POLETTO, N. Modelo para estimativa do potencial produtivo em trigo e cevada por meio do sensor GreenSeeker. *Engenharia Agrícola*. 2009. **29**(1), 101-112. https://doi.org/10.1590/S0100-69162009000100011

GROHS, D.S., BREDEMEIER, C., POLETTO, N. and MUNDSTOCK, C.M. Validação de modelo para predição do potencial produtivo de trigo com sensor óptico ativo. *Pesquisa Agropecuária Brasileira*. 2011. **46**(4), 446-449. https://doi.org/10.1590/S0100-204X2011000400015

HANSEN, P.M. and SCHJOERRING, J.K. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote sensing of environment.* 2003. **86**(4), 542-553. https://doi.org/10.1016/S0034-4257(03)00131-7

HAYKIN, S. *Redes neurais*: princípios e prática. Porto Alegre: Bookman, 2001.

JANA, R.B. and MOHANTY, B.P. A comparative study of multiple approaches to soil hydraulic parameter scaling applied at the hillslope scale. *Water Resources Research*. 2012, **48**(2), 2520. https://doi.org/10.1029/2010WR010185

JENSEN, J.R. *Introductory digital image processing:* a remote sensing perspective. Upper Saddle River: Prentice Hall, 1996.

KROSS, A., et al. Using Artificial Neural Networks and Remotely Sensed Data to Evaluate the Relative Importance of Variables for Prediction of Within-Field Corn and Soybean Yields. *Remote Sensing*. 2020, **12**(14), 2230. https://doi.org/10.3390/rs12142230

LU, J., et al. Evaluating different approaches to non-destructive nitrogen status diagnosis of rice using portable RapidSCAN active canopy sensor. *Scientific reports*. 2017, **7**(1), 1-10. https://doi.org/10.1038/s41598-017-14597-1

MOTOMIYA, A.V.A., et al. Índice de vegetação no algodoeiro sob diferentes doses de nitrogênio e regulador de crescimento. *Semina: Ciências Agrárias.* 2014, **35**(1), 169-177. https://doi.org/10.5433/1679-0359.2014v35n1p169

MUTANGA, O. and SKIDMORE, A.K. Integrating imaging spectroscopy and neural networks to map grass quality in the Kruger National Park, South Africa. *Remote sensing of environment*. 2004, **90**(1), 104-115. https://doi.org/10.1016/j.rse.2003.12.004

NAVALGUND, R.R. Remote sensing. *Resonance*. 2002, **7**(1), 37-46.

NOVO, E.M.L.M. *Sensoriamento Remoto:* princípios e aplicações. São Paulo: Blucher, 2008.

POVH, F.P., et al. Comportamento do NDVI obtido por sensor ótico ativo em cereais. *Pesquisa Agropecuária Brasileira*. 2008, **43**(80), 1075-1083. https://doi.org/10.1590/S0100-204X2008000800018

SANTOS, H.G., et al. *Sistema Brasileiro de Classificação de Solos*. 5th ed. Brasília: Embrapa, 2018.

SILVA, I.N., Spatti, D.H. and Flauzino, R.A. *Redes Neurais Artificiais:* para engenharias e ciências aplicadas. São Paulo: Artliber, 2016.

SOARES, F.C., Robaina, A.D., Peiter, M.X. and Russi, J.L. Predição da produtividade da cultura do milho utilizando rede neural artificial. *Ciência Rural.* 2015, **45**(11), 1987-1993. https://doi.org/10.1590/0103-8478cr20141524

SOARES, F.C., et al. Redes neurais artificiais na estimativa da retenção de água do solo. *Ciência Rural.* 2014, **44**(2), 293-300. https://doi.org/10.1590/S0103-84782014000200016

TOLEDO, G.L. and Ovalle, I.I. *Correlação e regressão*: Estatística Básica. 2nd ed. São Paulo: Atlas, 1995.

TRIMBLE, T. the W. the W.W. *Especificações técnicas.* Available from: http://www.geodata.eng.br/manuais/gps/trimble/R6.pdf

VIEIRA, T.G.C., LACERDA, W.S. and BOTELHO, T.G. Mapeamento de áreas cafeeiras utilizando redes neurais artificiais: Estudo de caso na região de Três Pontas, Minas Gerais. In: *Anais XIV Simpósio Brasileiro de Sensoriamento Remoto*,14, Natal, Brasil, 2009. Presented at the XIV Simpósio Brasileiro de Sensoriamento Remoto, INPE, 7947–7954.

ZHANG, Q. *Precision Agriculture Technology for Crop Farming*. Boca Raton: CRC Press, 2015.

Received: 6 July 2020 | **Accepted:** 21 October 2021 | **Published:** 31 March 2022