

Comparison of Methods for Harvest Prediction in ‘Gigante’ Cactus Pear

Bruno V. C. Guimarães¹, Sérgio L. R. Donato², Ignacio Aspiazú³, Alcinei M. Azevedo⁴ & Abner J. de Carvalho³

¹ Departamento de Ciências Agrárias, Instituto Federal do Amazonas, São Gabriel da Cachoeira, Amazonas, Brazil

² Setor de Agricultura, Instituto Federal Baiano, Guanambi, Bahia, Brazil

³ Departamento de Ciências Agrárias, Universidade Estadual de Montes Claros, Janaúba, Minas Gerais, Brazil

⁴ Instituto de Ciências Agrárias, Universidade Federal de Minas Gerais, Montes Claros, Minas Gerais, Brazil

Correspondence: Bruno V. C. Guimarães, Departamento de Ciências Agrárias, Instituto Federal Baiano, BR 307, km 03, Estrada do Aeroporto, São Gabriel da Cachoeira, Amazonas, Brazil. Tel: 55-779-9121-4608. E-mail: E-mail: bvincius20@yahoo.com.br

Received: May 14, 2019

Accepted: June 23, 2019

Online Published: August 31, 2019

doi:10.5539/jas.v11n14p216

URL: <https://doi.org/10.5539/jas.v11n14p216>

The research is financed by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Brazil)-Finance Code 001.

Abstract

Behavior analysis and plant expression are the answers the researcher needs to construct predictive models that minimize the effects of the uncertainties of field production. The objective of this study was to compare the simple and multiple linear regression methods and the artificial neural networks to allow the maximum security in the prediction of harvest in ‘Gigante’ cactus pear. The uniformity test was conducted at the Federal Institute of Bahia, Campus Guanambi, Bahia, Brazil, coordinates 14°13'30" S, 42°46'53" W and altitude of 525 m. At 930 days after planting, we evaluated 384 basic units, in which were measured the following variables: plant height (PH); cladode length (CL), width (CW) and thickness (CT); cladode number (CN); total cladode area (TCA); cladode area (CA) and cladode yield (Y). For the comparison between the artificial neural networks (ANN) and regression models (single and multiple-SLR and MLR), we considered the mean prediction error (MPE), the mean quadratic error (MQE), the mean square of deviation (MSD) and the coefficient of determination (R^2). The values estimated by the ANN 7-5-1 showed the best proximity to the data obtained in field conditions, followed by ANN 6-2-1, MLR (TCA and CT), SLR (TCA) and SLR (CN). In this way, the ANN models with the topologies 7-2-1 and 6-2-1, MLR with the variables total cladode area and cladode thickness and SLR with the isolated descriptors total cladode area and cladode number, explain 85.1; 81.5; 76.3; 74.09 and 65.87%, respectively, of the yield variation. The ANNs were more efficient at predicting the yield of the ‘Gigante’ cactus pear when compared to the simple and multiple linear regression models.

Keywords: model, experimental, *Opuntia ficus indica* Mill

1. Introduction

The Brazilian semi-arid region, circumscribed in the Caatinga biome, presents severe limits to plant production, mainly due to the low contents of available water in the soil for the plants (Albuquerque et al., 2018). However, even in this environment, unfavorable to plant growth and development, the cactus pear has emerged as a strategic resource in ruminant feeding (Ochoa et al., 2018) with significant levels of biomass production (Padilha Junior et al., 2016). For Amania et al. (2019), this productive potential is associated with several mechanisms of adaptation of the crop to the adverse conditions in which the species is normally conducted.

The agronomic performance of a crop is of paramount importance to those who are dedicated to rural entrepreneurship, especially in climatic conditions of high productive risk (Nalley et al., 2016). Thus, the analysis of the behavior and expression of the cultivated materials are the answers the researcher needs to

construct predictive models that minimize the effects of the productive uncertainties or the risk associated with the activity in the field (Guimarães et al., 2018).

The composition of models for harvest estimation with the ‘Gigante’ cactus pear is still sporadic in scientific literature. Although there are studies relating morphometric, morphogenic and production components in different cactus pear species (Padilha Junior et al., 2016) or even studies approaching morphological and their reflexes on yield (Silva et al., 2014), general literature lacks information that allows a comparison between the predictive tools related to cladode production in a practical and precise order, with direct application in the field.

Artificial neural networks (ANNs) stand out in predictive modeling (Fernandes et al., 2017), which is the supervised learning with multi-layer perceptron networks (MLP), trained with the back propagation algorithm, the most used in the field of prediction. The ANNs are architected in three structures, in which the predictor variables make up the first layer; the hidden layer relates the number of neurons to be scaled; in sequence, the output layer receives stimuli from the hidden layer and constructs the pattern that will be the response. With many applications, the use of ANNs has intensified in agricultural modeling (Aquino et al., 2016; Azevedo et al., 2017).

In the same sense, simple and multiple linear regression models have been extensively incorporated into agricultural prediction, such as in the estimation of irrigation depths (Vicente et al., 2015), reference evapotranspiration (Minuzzi et al., 2014), leaf area (Zeist et al., 2014) and yield (Bertolin et al., 2017). A comparison of these predictive models, ANNs and regressions, allows the researcher/producer to contrast the estimators in the field of modeling (Soares et al., 2014). With this, it is possible to identify a robust and consistent tool in agricultural prediction.

The objective of this study was to compare simple and multiple linear regression models and artificial neural networks to allow the maximum security in harvest prediction of ‘Gigante’ cactus pear.

2. Material and Methods

2.1 Experimental Characterization: Soil, Climate and Experimental Delimitation

The study was developed in the experimental field of the Baiano Federal Institute-IFBAIANO, Campus Guanambi, Bahia, Brazil, in a predominantly flat area, with soil classified as a Litolic Neosol, with the coordinates 14°13'30" S, 42°46'5" W and altitude of 525 m, with rainfall and average annual temperature of 670.2 mm and 25.9 °C, respectively.

The uniformity experiment with the ‘Gigante’ cactus pear was carried out with cladodes duly selected in the IFBAIANO matrix unit, being the preparation and curing of the seedlings in shaded conditions for 15 days. A total of 384 plants, spaced at 2.0 × 0.2 m, were designated as basic units (BUs) and circumscribed by a border with 116 plants, in a total area of 200 m².

2.2 Agronomic Characteristics Evaluated

The agronomic representation of the crop for data collection and analysis was obtained in the third production cycle, at 930 days after planting, at which the following predictive vegetative variables were measured: plant height (PH-m); cladode length (CL-cm), width (CW-cm) and thickness (CT-mm); cladode number (CN); total cladode area (TCA-cm²); cladode area (CA - cm²) and the yield variable to be predicted by the models-cladode yield (Y-t ha⁻¹).

2.3 Models for Harvest Prediction

In order to estimate the yield of ‘Gigante’ cactus pear, models of multi-layer perceptron (MLP) artificial neural networks (ANNs) with multiple neural structures were developed to describe crop yield. The most suitable models for prediction had three layers, the first with seven and six input neurons, a hidden layer with five and two neurons, and the output layer with one neuron, which represents the estimated variable.

The ANN models have as a characteristic the need of numerous trainings to express the prediction. Thus, to improve the networks, the function *mlp* of the RSNNS package with back propagation algorithm was applied, in which the synaptic weights, established during the training phase, are randomly generated in response to the reduction of the mean square error (MSE).

The training stages were sequenced in: 500 trainings, with activation in the hidden layer and output with the logistic and linear function, respectively; and 12 combinations of network architectures with 1, 2, 3, ... 9, 10, 20 and 30 neurons in the hidden layer, being each ANN architecture trained 100 times, selecting the best network by the lowest MSE value.

In order to optimize the prediction model with the maximum execution time efficiency and computational optimization, 1,000 new trainings were performed only with the ANN selected by the lower MSE, thus, avoiding numerous trainings for each network configuration.

The best ANN topology was developed using the back propagation algorithm to predict yield in ‘Gigante’ cactus pear by means of the phenotypic descriptors cladode area (CA), cladode length (CL), cladode width (CW), cladode thickness (CT), cladode number (CN), plant height (PH) and total cladode area (TCA). The ANNs were trained with the vegetative experimental data in order to reach the highest predictive capacity of fresh matter of the forage crop and the greater potential of generalization of the model for other applications.

In order to test the ANN efficiency, the phytotechnical data were divided into two groups, called training and validation, in the proportion of 80 and 20%, respectively. Thus, by the analysis of regression between estimated and observed yield in the validation sample, we evaluated the coefficient of determination (R^2) of the fitted model and the significance of the angular coefficient of the line by the t test, considering that an angular coefficient equal to one express the predictive capacity of the model by the high correspondence between the predicted and observed values.

In the simple and multiple regression analyzes, the models were fitted by the least squares procedure, with the regression models with simple effects and the multiple linear regression models with two or more predictor variables, respectively (Equation 1).

$$y_i = \beta_0 + \beta_{1 \times 1} + \beta_{2 \times 2} + \dots + \beta_{k \times k} + e_i \quad (1)$$

where, y_i refers to the fresh matter of cladodes based on the yield-related variables (x_1, \dots, x_k) CA, CL, CW, CT, CN, PH and TCA. The e_i is the error associated with i^{th} -observation, with normal and independent distribution, the constant β_0 is the intercession point of the model and $\beta_1 + \dots + \beta_{k \times k}$ represent the coefficients.

Based on the presented model, the stepwise method (Draper & Smith, 1981) was used to select the most relevant variables. In addition to this procedure, the selection criterion considered only the significant variables at the 5% significance level by Student’s t test. The Akaike Information Criterion-AIC-Equation 2 (Akaike, 1974) was also considered as an adjustment quality estimator to select simple and multiple linear regression models, defined by:

$$AIC = -2 \ln(L_p) + 2p \quad (2)$$

where, L_p defines the maximum likelihood function of the estimated model, and p is the number of parameters associated with the model.

For the comparison between the ANNs and regression models, the mean prediction error (MPE-Equation 3), the mean quadratic error (MQE), the mean square of the deviations (MSD-Equation 4), the fitted coefficient of determination and the coefficient of determination (R^2 -Equation 5),

$$EMP(\%) = \sum_i^n \frac{(X_{obs} - X_{pred}) \times 100 / X_{obs}}{n} \quad (3)$$

$$QMD = \sum_i^n \frac{(X_{obs} - X_{pred})^2}{n} \quad (4)$$

where, \sum_i^n is the sum from i to n ; x_{obs} is the fresh matter of cladode measured after harvest; x_{pred} is the matter estimated by the ANN and MRL models and “ n ” is the number of observations.

$$R^2 = \frac{SS_{Reg}}{TSS} = 1 - \frac{RSS}{TSS} \quad (5)$$

where, $0 \leq R^2 \leq 1$.

2.4 Statistical Analysis

The ANN and regression models were generated using the R software (R Development Core Team, 2019).

3. Results and Discussion

The vegetative variables plant height and area, width, length and thickness of the cladodes, although showing significant coefficients, are not recommended to compose the models of yield prediction due to the low correlation and, consequently, limited capacity to explain the behavior and the expression of yield, as indicated by R^2 and R^2a (Table 1). By the simple linear regression model, the isolated variable total area of cladodes allows to predict the yield with adequate values of R^2 and R^2a , being this model also selected by the lowest estimate of AIC.

Table 1. Components of the simple linear regression model for prediction of ‘Gigante’ cactus pear yield by vegetative variables

Constant	Variables	Coefficients	AIC	r	R ²	R ² a
-21.890	TCA	252.876	4445.291	0.8608	0.7409	0.7402
-33.2014	CN	16.2426	4551.134	0.8116	0.6587	0.6578
-120.04	PH	362.38	4840.924	0.5235	0.274	0.2721
-103.0273	CA	1.3324	4909.228	0.3643	0.1327	0.1304
-349.974	CW	45.195	4915.467	0.3442	0.1185	0.1161
-442.955	CL	25.227	4917.611	0.3369	0.1135	0.1112
178.095	CT	9.440	4939.758	0.2468	0.06089	0.05843

Note. TCA: total cladode area; CN: cladode number; PH: plant height; CA: area, CW: width, CL: length and CT: thickness of the cladodes; AIC: Akaike Information Criterion; r: correlation coefficient; R²: coefficient of determination; R²a: adjusted coefficient of determination.

The predictive descriptors cladode number, plant height, area, width and length of cladodes were not included in the equation by multiple linear regression analysis, as this method only considers the significant variables. Soares et al. (2014) argue that the variables discarded by the method possibly exert little influence on the predicted variable. Consequently, for the multiple linear regression model, only the vegetative variables were selected for total cladode area and cladode thickness (Table 2).

Table 2. Components of the simple linear regression model for prediction of ‘Gigante’ cactus pear yield by vegetative variables

Constant	Coefficients								R ²	R ² a
	PH	CN	TCA	CA	CT	CL	CW	AIC		
-93.7883	-	-	247.7903	-	5.6736	-	-	4413.71	0.763	0.761

Note. TCA: total cladode area; CN: cladode number; PH: plant height; CA: area, CW: width, CL: length and CT: thickness of the cladodes; AIC: Akaike Information Criterion; R²: coefficient of determination; R²a: adjusted coefficient of determination.

The level of significance attributed to the predictor variables allows distinguishing which descriptor has the greatest influence on yield. Thus, the prediction studies aim to quantify the effect that one or more vegetative characters can cause on a response variable. In this work, the R²a for the models selected with simple linear regression ranged from 0.6578 to 0.7402 (Table 1). The multiple linear equation added only 2% in the explanatory capacity of the model, with the R²a of 0.761 (Table 2).

Soares et al. (2014) fitted regression models with similar R² when estimating the matter of the bunches in ‘Tropical’ bananas. The coefficients of determination of the models presented in this study, such as those related to the estimation of production in other crops (Leal et al., 2015; Kaytez et al., 2015; Dornelles et al., 2018) can be considered of low magnitude, since the predictive descriptors explained in a limited way the performance of the predicted variable

After intense training for the composition of the ANNs, two architectures with the greatest potential to predict cactus pear yield were determined based on the coefficient of determination, mean square error and mean square of the deviations. The models 7-5-1 and 6-2-1 (Table 3) were selected using the criteria of relevance and greater predictive capacity. The first one had all the measured variables in the input layer (CA, CL, CW, CT, CN, PH, TCA) and five neurons in the middle layer, whereas the second estimator did not include the phenotypic descriptor CA and only had two neurons in the hidden layer.

Table 3. Statistical parameters for artificial neural network (ANN) architecture with seven and six variables in the input layer

ANNs for sampling data	7-1-1	7-2-1	7-3-1	7-4-1	7-5-1	7-6-1	7-7-1	7-8-1	7-9-1	7-10-1	7-20-1	7-30-1
<i>Network architecture with seven variables in the input layer</i>												
R ²	0.687	0.723	0.710	0.766	0.851	0.749	0.769	0.699	0.809	0.791	0.751	0.762
MSE	0.011	0.009	0.007	0.010	0.004	0.011	0.010	0.008	0.007	0.006	0.010	0.006
MSD	3.05	2.919	2.467	2.818	1.977	3.029	2.773	2.499	2.523	2.285	2.662	2.158
<i>Network architecture with six variables in the input layer</i>												
	6-1-1	6-2-1	6-3-1	6-4-1	6-5-1	6-6-1	6-7-1	6-8-1	6-9-1	6-10-1	6-20-1	6-30-1
R ²	0.798	0.815	0.763	0.769	0.759	0.725	0.746	0.711	0.780	0.701	0.711	0.788
MSE	0.007	0.006	0.005	0.009	0.010	0.008	0.008	0.010	0.007	0.010	0.008	0.007
MSD	3.555	2.858	2.928	3.561	3.728	3.369	3.338	3.655	3.095	3.820	3.326	3.167

Note. R²: coefficient of determination; MSE: Mean squared error; MSD: mean square of the deviations.

In addition to representing all tested network architectures, Table 3 presents the statistical parameters associated with the predictive models. Thus, the R², MSE and MSD markers attest to the high predictive capacity of the 7-5-1 and 6-2-1 models, with 0.851; 0.004 and 1.983 and 0.815; 0.006 and 2.858 (Table 3), respectively. The R² values indicate that the vegetative variables used allow to explain in 85.1 and 81.5% the productive variation of the cactus pear with the architectures 7-5-1 and 6-2-1, respectively.

Soares et al. (2014) and Leal et al. (2015) fitted the best models to estimate yield with the 10-10-1 and 4-4-1 network formats for banana and corn crops, in that order. Similarly to the present study, the researchers emphasized that the consecutive increase in the number of neurons in the hidden layer does not represent a tendency to the best adjustment of the network, as the intense training through the combination of the synaptic weights defines the most appropriate architecture to estimate the dependent variable.

In order to predict the yield of the ‘Gigante’ cactus pear, the improvement of the ANNs considered an error tolerance of 0.0001, arbitrated in 500 trainings, with a learning rate of 0.1. These values, previously established, were defined in the learning process to ensure the maximum efficiency of the ANNs, since smaller values would overload the analyses without positive effect on the results (Soares et al., 2014).

Figure 1 represents the best ANN topology developed by the back propagation algorithm to predict yield in ‘Gigante’ cactus pear, using the phenotypic descriptors CA, CL, CW, CT, CN, PH and TCA (Figure 1A) and the predictive model without the variable CA (Figure 1B).

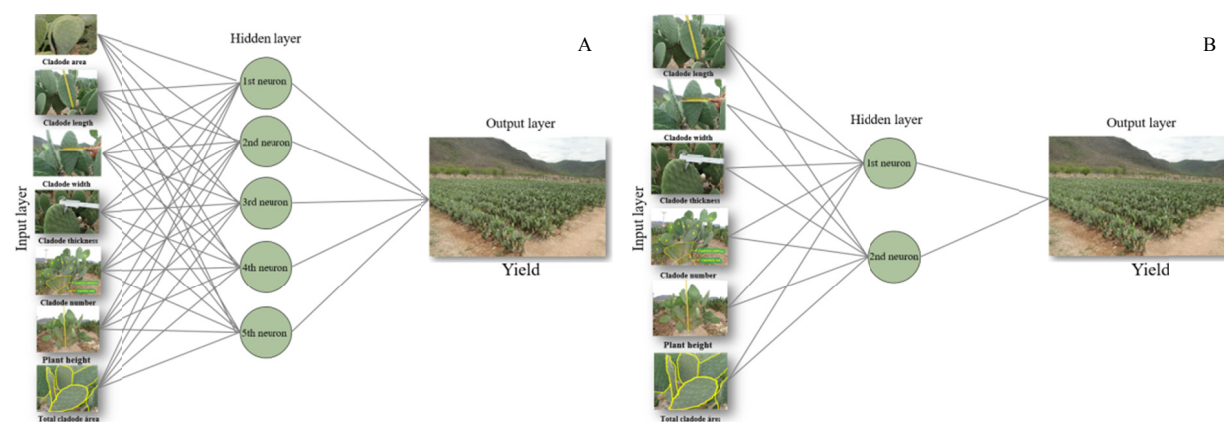


Figure 1. ANN architecture for yield prediction in ‘Gigante’ cactus pear using the phenotypic descriptors cladode area (CA), cladode length (CL), cladode width (CW), cladode thickness (CT), cladode number (CN), plant height (PH) and total cladode area (TCA) (A) and the predictive model without the descriptor CA (B)

After an intense training with the interconnections of the network neurons from the vegetative variables measured in the third production cycle, ANNs 7-5-1 and 6-2-1 made it possible to accurately visualize the effect of the network on the output layer-cactus pear yield. Thus, these models make it possible to predict the practical

reality of the planned areas with this crop by means of the vegetative variables inserted in the input layer, as in Figure 1.

Figure 2 graphically compares the values estimated by the SLR, MLR and ANN models in contrast to the original yield data measured in the field. The estimators-SLR with the isolated variables ATC and NC; MLR with the conjugated variables ATC and CT; and ANNs with the architectures 7-5-1 and 6-2-1 were associated with the red, purple, yellow, blue and green lines, in this order; yield is represented by the black line with the field data. Thus, the closer the model lines were to the black line, the greater the accuracy and reliability of the yield forecast characteristic.

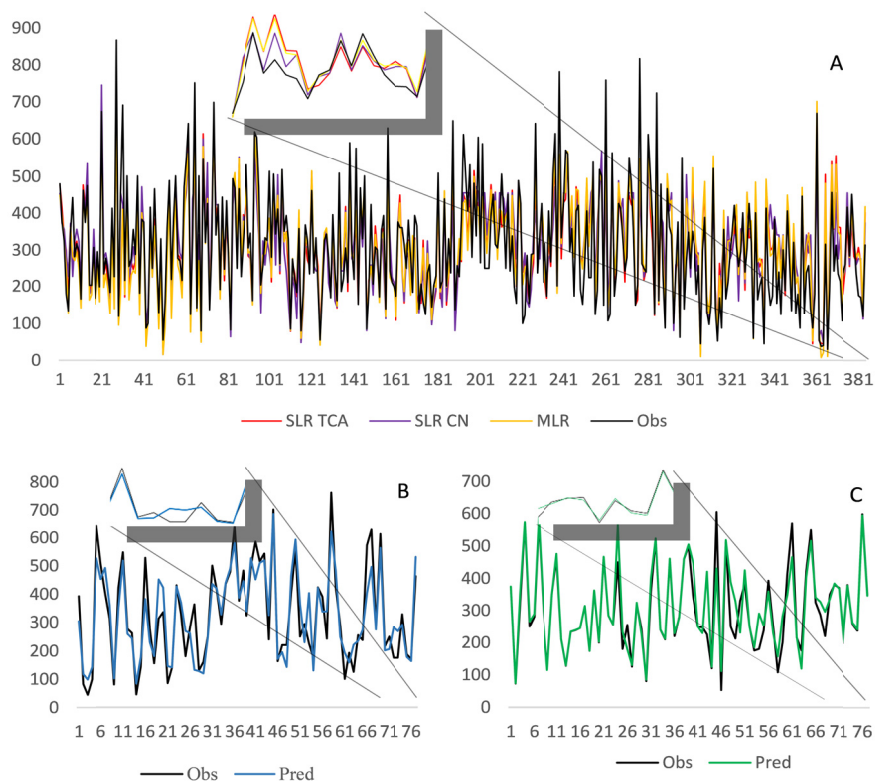


Figure 2. Graphical representation of the observed data (black), data estimated by the SLR with the variables TCA (red) and CN (purple), data predicted by MLR with the variables TCA and CT (yellow) and data predicted by ANN 6-2-1 (B-blue) and 7-5-1 (C-green) for ‘Gigante’ cactus pear yield prediction. We used 384 and 77 plants for the regression and ANN models, respectively. The zoom cuts contemplated ten plants for all models

The values estimated by the ANN with the architecture 7-5-1 (Figure 2C) presented the closest proximity to the data measured in the field, followed by the values estimated by ANN 6-2-1, MLR (TCA and CT), SLR (TCA) and SLR (CN). These graphical assumptions are corroborated by the coefficients of determination of the models, in which the behavior of the response variable is explained by the predictor variables. In these terms, the ANNs with the topologies 7-2-1 and 6-2-1, the MLR with the variables TCA and CT, and the SLR with the isolated variables ATC and NC explain 85.1; 81.5; 76.3; 74.09 and 65.87%, respectively, of ‘Gigante’ cactus pear yield variation (Table 4).

Table 4. Comparison between simple and multiple linear regression models and artificial neural networks for yield estimation in ‘Gigante’ cactus pear

Characteristic	SLR (CN)		SLR (TCA)		MLR (TCA & CT)		ANN 6-2-1		ANN 7-5-1	
	R ²	MPE	R ²	MPE	R ²	MPE	R ²	MPE	R ²	MPE
Yield	0.659	2.11	0.741	1.94	0.763	1.74	0.815	0.198	0.851	0.186

Note. SLR: simple linear regression; MLR: multiple linear regression; ANN: artificial neural network; R²: coefficient of determination; MPE (%): mean prediction error.

The combination of the values predicted by the models and the original data made it possible to obtain the R^2 and the calculation of the mean error associated with the prediction (Table 4). The highest and the lowest mean errors of the prediction were associated with the SLR (NC) and ANN 7-5-1 models, respectively. In a complementary way, Figure 3 represents the accuracy of the models studied through the behavior of the respective lines, and the best fit between the observed data and the estimated values were defined with ANNs (Figure 3C). Fernandes et al. (2017) add that the superiority of this method is due to the sensitivity in the inflection rates of the progression curves, which reflects the greater predictive capacity.

For Soares et al. (2014), the prediction models with ANNs showed great potential in describing the matter of the bunch in 'Tropical' bananas, with MPE of 1.40, while multiple linear regression was fitted with MPE of 6.52. These results were similar to the present study, with the lowest prediction errors and the highest R^2 values associated to the ANNs. Hence, these can be considered as a useful tool to delineate the plant expression and behavior.

Due to the safety and efficiency of the technique in decision making, the ANNs have stood out in several lines of prediction because of the generalization capacity of the trained models in up to 100%. This predictive architecture was selected by Campos and Garcia (2019) in the format 26-12-1 and 0.0001, referring, respectively, to the number of neurons in the input layer, the hidden layer, the output layer and the mean square error. Fernandes et al. (2017) concluded that ANN modelling, although with R^2 equal to 0.61, was more effective in estimating sugarcane yield than official surveys, anticipating the harvest forecast in three months.

One of the main applications of ANNs is anchored in the prediction of phenomena (Campos & Garcia, 2019), a circumstance also related to regression models (Bertolin et al., 2017). Thus, in view of this convergence or functional similarity, some studies were developed with the objective of comparing the efficiency of ANN models with regression techniques. In the evaluations of these prediction methods in several crops such as maize (Leal et al., 2015), banana (Soares et al., 2014) and rice (Giordano et al., 2010), the authors demonstrated a higher predictive quality for ANNs in relation to the regression equations.

In addition, other studies (Fernandes et al., 2017; Campos et al., 2017; Dornelles et al., 2018) described the ANNs as robust and efficient tools with a remarkable predictive capacity, in which the historical patterns of a given data set can be projected into refined trending lines, with a view to solving problems and providing responses to decision making.

Leal et al. (2015) argue that, although the ANNs present higher computational costs, with the greater demand for the construction of networks and the need to be trained countless times at each validation step, the use of this technique, based on the clay content variables, cation exchange capacity, soil organic matter and base saturation, allows better adjustments to estimate grain yield in relation to regression estimators.

Soares et al. (2013) ratified the accuracy and efficiency of the ANN computational model in estimating yield in 'Tropical' bananas. Faced with the quality of the fitted model, the same authors compared ANN estimators and multiple linear regression for prediction of the same variable and found differences between the methods, similarly to the present study, with R^2 of 0.91 and 0.71, in this order, for the tested models (Soares et al., 2014).

Dornelles et al. (2018) simulated models for predicting oat grain yield through artificial intelligence and traditional polynomial regression analyzes, identifying higher performance and predictive quality in fitting with the use of artificial neural networks of multiple layers. Thus, the artificial model, as it presents a smaller error associated to the prediction, allows us to construct strategies to optimize the agricultural resources and to make feasible marketing plans with greater security (Leal et al., 2015).

Campos et al. (2017) have identified similarities in predictive efficiency between ANN models and the regression equations traditionally employed by forest companies. However, the models with the ANNs, similarly to the present study, presented higher coefficient of determination and smaller value in the square root of the mean error. The comparison between prediction tools in agriculture has been growing significantly, especially for providing the researcher/producer tools to ensure maximum efficiency in decision making (Arruda et al., 2013; Soares et al., 2014; Leal et al., 2015; Dornelles et al., 2018).

The development of high-efficiency models obtained from the original field conditions favors safety when predicting 'Gigante' cactus pear yield (Guimarães et al., 2018), which contributes to the success of rural planning, especially as a support to the producer who needs to define in advance the quantity of animals to be fed with the produced biomass or even the volume to be marketed (Marques et al., 2017).

Thus, the solution of problems through artificial intelligence, using artificial neural networks, is quite significant and promising to substantiate decisions in agriculture (Soares et al., 2014; Dornelles et al., 2018), promoting

models predictive models with superior performance compared to validated conventional tools (Arruda et al., 2013; Leal et al., 2015; Kaytez et al., 2015).

4. Conclusions

The ANNs allow the development of more efficient models for the prediction of ‘Gigante’ cactus pear yield in comparison to the simple and multiple linear regression models, using the vegetative variables cladode area, cladode length, cladode width, cladode thickness, cladode number, plant height and total cladode area.

Acknowledgements

To the Federal Institutes: Instituto Federal do Amazonas (IFAM) and Instituto Federal Baiano (IF-Baiano) and to the Universidade Estadual de Montes Claros (UNIMONTES). The present study was conducted with support from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Brazil)-Finance Code 001.

References

- Albuquerque, J. G., Aquino, J. S., Albuquerque, J. G., Farias, T. G. S., Buendía, H. B. E., Molina, E. B., & Azoubel, P. M. (2018). Consumer perception and use of nopal (*Opuntia ficus-indica*): A cross-cultural study between Mexico and Brazil. *Food Research International*, *1*, 1-41. <https://doi.org/10.1016/j.foodres.2018.08.036>
- Amania, E., Marwa, L., Hichem, B. S., Amel, S. H., & Ghada, B. (2019). Morphological variability of prickly pear cultivars (*Opuntia* spp.) established in ex-situ collection in Tunisia. *Scientia Horticulturae*, *248*(5), 163-175. <https://doi.org/10.1016/j.scienta.2019.01.004>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716-723. <https://doi.org/10.1109/TAC.1974.1100705>
- Aquino, C. F., Salomão, L. C. C., & Azevedo, A. M. (2016). Fenotipagem de alta eficiência para vitamina A em banana utilizando redes neurais artificiais e dados colorimétricos. *Bragantia*, *75*(1), 268-274. <https://doi.org/10.1590/1678-4499.467>
- Arruda, G. P. de, Demattê, J. A. M., & Chagas, C. da S. (2013). Mapeamento digital de solos por redes neurais artificiais com base na relação solo-paisagem. *Revista Brasileira de Ciência do Solo*, *37*(2), 327-338. <https://doi.org/10.1590/S0100-06832013000200004>
- Azevedo, A. M., Andrade Júnior, V. C., Sousa Júnior, A. S., Santos, A. A., Cruz, C. D., Pereira, S. L., & Oliveira, A. J. M. (2017). Eficiência da estimação da área foliar de couve por meio de redes neurais artificiais. *Horticultura Brasileira*, *35*(1), 14-19. <https://doi.org/10.1590/s0102-053620170103>
- Bertolin, N. de O., Filgueiras, R., Venancio, L. P., & Mantovani, E. C. (2017). Predição da produtividade de milho irrigado com auxílio de imagens de satélite. *Revista Brasileira de Agricultura Irrigada*, *11*(4), 1627-1638. <https://doi.org/10.7127/rbai.v11n400567>
- Campos, B. P. F., da Silva, G. F., Binoti, D. H. B., de Mendonça, A. R., & Leite, H. G. (2017). Descrição do perfil do tronco de árvores em plantios de diferentes espécies por meio de redes neurais artificiais. *Pesquisa Florestal Brasileira*, *37*(90), 99-107. <https://doi.org/10.4336/2017.pfb.37.90.1181>
- Campos, L. M. L., de & Garcia, J. C. P. (2019). Redes neurais apoiando a tomada de decisão na análise de crédito bancário e detecção do câncer de mama. *Revista Gestão & Tecnologia*, *19*(1), 90-112. <https://doi.org/10.20397/2177-6652/2019.v19i1.1451>
- Dornelles, E. F., Kraissig, A. R., Silva, J. A. G. da, Sawicki, S., Frantz, R., & Carbonera, R. (2018). Artificial intelligence in seeding density optimization and yield simulation for oat. *Revista Brasileira de Engenharia Agrícola e Ambiental*, *22*(3), 183-188. <https://doi.org/10.1590/1807-1929/agriambi.v22n3p183-188>
- Draper, N. R., & Smith, H. (1981). *Applied Regression Analysis* (2nd ed., p. 709). J. Wiley, New York.
- Fernandes, J. L., Ebecken, N. F. F., & Esquerdo, J. C. D. M. (2017). Sugarcane yield prediction in Brazil using NDVI time series and neural networks ensemble. *International Journal of Remote Sensing*, *38*(1), 4631-4644. <https://doi.org/10.1080/01431161.2017.1325531>
- Giordano, P. C., Martinês, H. D., Iglesias, A. A., Beccaria, A. J., & Goicoechea, H. C. (2010). Application of response surface methodology and artificial neural networks for optimization of recombinant *Oryza sativa* non-symbiotic hemoglobin 1 production by *Escherichia coli* in medium containing byproduct glycerol. *Bioresource Technology*, *101*(1), 7537-7544. <https://doi.org/10.1016/j.biortech.2014.01.021>

- Guimarães, B. V. C., Donato, S. L. R., Azevedo, A. M., Aspiazú, I., & Silva Junior, A. A. (2018). Prediction of 'Gigante' cactus pear yield by morphological characters and artificial neural networks. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 22(5), 315-319. <https://doi.org/10.1590/1807-1929/agriambi.v22n5p315-319>
- Kaytez, F., Taplamacioglu, C., Cam, E., & Hardalac, F. (2015). Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. *International Journal of Electrical Power & Energy Systems*, 67(1), 431-438. <https://doi.org/10.1016/j.ijepes.2014.12.036>
- Leal, A. J. F., Miguel, E. P., Baio, F. H. R., Neves, D. de C. N., & Leal, U. A. S. (2015). Redes neurais artificiais na predição da produtividade de milho e definição de sítios de manejo diferenciado por meio de atributos do solo. *Bragantia*, 74(4), 436-444. <https://doi.org/10.1590/1678-4499.0140>
- Marques, O. F. C., Gomes, L. S. P., Mourthé, M. H. F., Braz, T. G. dos S., & Pires Neto, O. de S. (2017). Palma forrageira: Cultivo e utilização na alimentação de bovinos. *Caderno de Ciências Agrárias*, 9(1), 75-93.
- Minuzzi, R. B., Ribeiro, A. J., Silva, D. O. da, & Kuneski, A. C. (2014). Estimativa da evapotranspiração de referência diária por penman-monteith FAO com dados de temperatura do ar para Santa Catarina. *Irriga*, 19(4), 548-558. <https://doi.org/10.15809/irriga.2014v19n4p548>
- Nalley, L., Tackb, J., Barkleyc, A., Jagadishd, K., & Bryee, K. (2016). Quantifying the Agronomic and Economic Performance of Hybrid and Conventional Rice Varieties. *Crop Economics, Production & Management*, 108(4), 1514-1523. <https://doi.org/10.2134/agronj2015.0526>
- Ochoa, G. O., Caratachea, A. J., Sánchez, R. E. P., Flores, H. E. M., Cordova, J. E., & Rodríguez, R. O. (2018). Effect of lactating sows' diet supplemented with cactus (*Opuntia ficus-indica*) on feed intake and reproductive and productive post-weaning performances. *Tropical Animal Health and Production*, 50(1), 1671-1682. <https://doi.org/10.1007/s11250-018-1611-x>
- Padilha Junior, M. C., Donato, S. L. R., Silva, J. A. S., Donato, P. E. R., & Souza, E. S. S. (2016). Características morfométricas e rendimento da palma forrageira 'Gigante' sob diferentes adubações e configurações de plantio. *Revista Verde de Agroecologia e Desenvolvimento Sustentável*, 11(1), 67-72. <https://doi.org/10.18378/rvads.v11i1.3710>
- R Development Core Team. (2012). *The R Project for Statistical Computing*. Retrieved from <https://www.r-project.org>
- Soares, J. D. R., Pasqual, M., Lacerda, W. S., Silva, S. O., & Donato, S. L. R. (2013). Utilization of artificial neural networks in the prediction of the bunches' weight in banana plants. *Scientia Horticulturae*, 155(1), 24-29. <https://doi.org/10.1016/j.scienta.2013.01.026>
- Soares, J. D. R., Pasqual, M., Lacerda, W. S., Silva, S. O., & Donato, S. L. R. (2014). Comparison of techniques used in the prediction of yield in banana plants. *Scientia Horticulturae*, 167(1), 84-90. <https://doi.org/10.1016/j.scienta.2013.12.012>
- Silva, T. G. F., Miranda, K. R., Santos, D. C., Queiroz, M. G., Silva, M. C., Cruz Neto, J. F. E., & Araújo, J. E. M. (2014). Área do cladódio de clones de palma forrageira: modelagem, análise e aplicabilidade. *Agrária. Revista Brasileira de Ciências Agrárias*, 9(4), 633-641. <https://doi.org/10.5039/agraria.v9i4a4553>
- Vicente, M. R., Mantovani, E. C., Fernandes, A. L. T., Delazari, F. T., & Figueredo, E. M. (2015). Efeito de diferentes lâminas de irrigação nas variáveis de desenvolvimento e produção do cafeeiro irrigado por pivô central. *Irriga*, 20, 528-543. <https://doi.org/10.1016/j.scienta.2013.12.012>
- Zeist, A. R., Oliveira, J. R. F. de, Lima Filho, R. B. de, Silva, M. L. de S., & Resende, J. T. V. de (2014). Comparação de métodos de estimativa de área foliar em morangueiro. *Pesquisa Agropecuária Gaúcha*, 20(3), 33-41.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).