



DOI: <http://dx.doi.org/10.1590/1807-1929/agriambi.v24n11p721-727>

Regression models for productivity prediction in cactus pear cv. Gigante

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

¹ Instituto Federal do Amazonas/Departamento de Ciências Agrárias. São Gabriel da Cachoeira, AM, Brasil. E-mail: bviniicius20@yahoo.com.br (Corresponding author) - ORCID: 0000-0003-2585-8794

² Instituto Federal Baiano/Setor de Agricultura. Guanambi, BA, Brasil. E-mail: sergio.donato@ifbaiano.edu.br - ORCID: 0000-0002-7719-4662

³ Universidade Estadual de Montes Claros/Departamento de Ciências Agrárias. Janaúba, MG, Brasil. E-mail: ignacio.aspiazu@unimontes.br - ORCID: 0000-0002-0042-3324; abner.carvalho@unimontes.br - ORCID: 0000-0002-6644-5307

⁴ Universidade Federal de Minas Gerais/Instituto de Ciências Agrárias. Montes Claros, MG, Brasil. E-mail: alcineimistico@hotmail.com - ORCID: 0000-0001-5196-0851

ABSTRACT: The understanding of plant behavior and its reflexes on yield is essential for rural planning; thus, the biomathematical models are promising in the yield prediction of cactus pear cv. Gigante. This study aimed to adjust, through simple and multiple regression analysis, models for predicting the yield of cactus pear cv. Gigante. The study, using homogeneous treatments, was developed at the Instituto Federal Baiano, Campus of Guanambi, Bahia, Brazil. Data were collected in an area consisting of 384 basic units (plants), in which the yield, defined as a dependent variable, and the predictor variables: plant height (PH), cladode length (CL), cladode width (CW), and cladode thickness (CT), number of cladodes (NC), cladode area (CA), and total cladode area (TCA) were evaluated. Simple linear regression models, multiple regression models only with simple effects for the explanatory variables, and the multiple regression models considering the simple and quadratic effects, and all its possible interactions were adjusted. From this last model, a reduced model was obtained by discarding the less relevant effects, using the Stepwise methodology. The use of the vegetative traits, TCA, NC, CA, CL, CT, and CW, through the adoption of multiple linear regression, quadratic interaction or just the variable TCA by the use of simple linear regression, allows the yield prediction of cactus pear, with adjusted R^2 of 0.82, 0.76, and 0.74, respectively.

Key words: *Opuntia* sp., modeling, estimation, yield

Modelos de regressão para predição de produtividade em palma forrageira cv. Gigante

RESUMO: O entendimento sobre o comportamento vegetal e seus reflexos sobre a produtividade é essencial ao planejamento rural, com isso, os modelos biomatemáticos são promissores na predição da produtividade da palma forrageira cv. Gigante. Objetivou-se com este estudo ajustar por meio de análises de regressão simples e múltipla modelos para predição da produtividade da palma forrageira cv. Gigante. O estudo, em formato de homogeneidade de tratamentos, foi desenvolvido no Instituto Federal Baiano, Campus Guanambi, Bahia, Brasil. Os dados foram coletados em área constituída de 384 unidades básicas, nas quais se mensuraram a produtividade, definida como variável dependente, e as variáveis predictoras: altura da planta (PH), comprimento (CL), largura (CW) e espessura do cladódio (CT), número de cladódio (NC), área do cladódio (CA), e área total do cladódio (TCA). Foram ajustados modelos de regressão linear simples, modelos de regressão múltipla com efeitos simples apenas para as variáveis explicativas e modelos de regressão múltipla, considerando tanto os efeitos simples, quadráticos e todas as suas interações possíveis. A partir deste último modelo citado, foi obtido um modelo reduzido pelo descarte dos efeitos menos relevantes, por meio da metodologia Stepwise. O uso das características vegetativas TCA, NC, CA, CL, CT e CW, por meio de adoção da regressão linear múltipla, interação quadrática ou somente a variável TCA pelo emprego da regressão linear simples, permite a predição da produtividade da palma forrageira, com R^2 ajustado de 0,82, 0,76 e 0,74, respectivamente.

Palavras-chave: *Opuntia* sp., modelagem, estimativa, rendimento



INTRODUCTION

The cactus pear cv. Gigante (*Opuntia ficus-indica* Mill.) presents excellent forms of adaptations to the semiarid ecosystem, mainly due to the photosynthetic process CAM (Crassulacean Acid Metabolism) characterized by stomatal opening and CO₂ capture at night (Taiz et al., 2017), and with efficient mechanisms of the water use (Silva et al., 2015).

Due to the high nutritional, energy, and water value, this forage stands out as a strategic food source in the nutrition of ruminants. Likewise, besides the potential to meet the needs of the herd, in balanced diets, the species assumes singular importance in the period of food scarcity and water restriction (Marques et al., 2017).

However, the success of agricultural activity goes beyond production. In this context, proper planning is essential because it allows the producer a tool to estimate production by non-destructive morphometric measures (Guimarães et al., 2013; 2018; 2019). Thus, organize a technical reserve to ensure the raw material supply to the animals continuously and safely, especially in advance of the dry season.

In the search for understanding about which vegetative descriptors are most associated with the production, as well as the possibility of using these to predict yield (Guimarães et al., 2014), aiming at defining the number of animals to be fed or biomass volume to be commercialized, the use of simple linear regression (SLR) (Bertolin et al., 2017), multiple linear regression (MLR) (Soares et al., 2014; Mantai et al., 2015) and the polynomial and quadratic regression models (Amaral et al., 2017) have been used as a reliable tool.

Given the above, the analysis of plant behavior and its reflexes on productivity is essential for rural planning. Thus, biomathematical models are promising in the prediction of crop yield. Therefore, this study aimed to adjust, through simple and multiple regression analysis, models for predicting the yield of cactus pear cv. Gigante.

MATERIAL AND METHODS

The study was carried out at Instituto Federal Baiano, Campus of Guanambi, Bahia, Brazil, between 2009 and 2011, at geographical coordinates, 14°13'30" S, 42°46'53" W and altitude of 525 m. The soil was classified as *Entisols Lithic*. The average annual precipitation and temperature are 670.2 mm and 25.9 °C, respectively (CODEVASF, 2018).

The study followed the format of treatment homogeneity or uniformity trial, in which the entire area implanted with the cactus pear cv. Gigante was submitted to the same agronomic conditions and evaluated at 930 days after planting (DAP) in the third production cycle.

The useful planting area was composed of eight central rows, with 48 plants per row, making 384 basic units (plants). The fresh mass yield of the cladodes (Prod, t ha⁻¹), considered the response variable, was determined in the third production cycle. Also, the following predictor variables were evaluated, plant height (PH, cm); cladode length (CL, cm); cladode width (CW, cm), measured using a graduated measuring tape; cladode thickness (CT, mm), defined by the caliper measuring

in the central part of the cladode; the number of cladodes (NC, n°), direct count; cladode area (CA, cm²), and total cladode area (TCA, m²), which were estimated by Eqs. 1 and 2, respectively, according to models adopted by Donato et al. (2014) and Padilha Junior et al. (2016).

$$CA = CL \cdot CW \cdot 0.693 \quad (1)$$

$$TCA = \left[\frac{(CA \cdot NC)}{10,000} \right] \cdot 2 \quad (2)$$

where:

- CA - cladode area, cm²;
- CL - cladode length, cm;
- CW - cladode width, cm;
- TCA - total cladode area, m²; and,
- NC - number of cladodes.

Through Pearson's correlation, associations between the morphological variables analyzed were evaluated. In the sequence, the simple linear regression models (Eqs. 3 and 4), the multiple regression models only with main effects for the explanatory variables (Eqs. 5 and 6), and the multiple regression models considering simple and quadratic effects and all its possible interactions (Eq. 7) were adjusted by the methods of least squares.

$$\text{Prod}_i = \beta_0 + \beta_1 \text{TCA}_i + e_i \quad (3)$$

$$\text{Prod}_i = \beta_0 + \beta_1 \text{NC}_i + e_i \quad (4)$$

$$\begin{aligned} \text{Prod}_i = & \beta_0 + \beta_1 \text{PH}_i + \beta_2 \text{TCA}_i + \beta_3 \text{NC}_i + \\ & + \beta_4 \text{CA}_i + \beta_5 \text{CL}_i + \beta_6 \text{CT}_i + \beta_7 \text{CW}_i + e_i \end{aligned} \quad (5)$$

$$\begin{aligned} \text{Prod}_i = & \beta_0 + \beta_1 \text{TCA}_i + \beta_2 \text{CA}_i + \beta_3 \text{CL}_i + \\ & + \beta_4 \text{CT}_i + \beta_5 \text{CW}_i + e_i \end{aligned} \quad (6)$$

$$\begin{aligned} \text{Prod}_i = & \beta_0 + \beta_1 \text{CW}_i + \beta_2 \text{CT}_i + \beta_3 \text{TCA}_i + \\ & + \beta_4 \text{CW}_i^2 - \beta_5 \text{CT}_i^2 + \beta_6 \text{TCA}_i^2 + \beta_7 \text{CL}_i \text{CT}_i + \\ & + \beta_8 \text{CW}_i \text{TCA}_i + \beta_9 \text{CW}_i \text{TCA}_i^2 + \beta_{10} \text{PH}_i \text{NC}_i + \\ & + \beta_{11} \text{PH}_i \text{CL}_i^2 + \beta_{12} \text{PH}_i \text{CT}_i^2 + \beta_{13} \text{PH}_i \text{NC}_i^2 + \\ & + \beta_{14} \text{NC}_i \text{PH}_i^2 + \beta_{15} \text{TCA}_i \text{CW}_i^2 + \beta_{16} \text{TCA}_i \text{PH}_i^2 + \\ & + \beta_{17} \text{CW}_i^2 \text{CT}_i^2 + \beta_{18} \text{CW}_i^2 \text{TCA}_i^2 + \beta_{19} \text{CT}_i^2 \text{PH}_i^2 + \\ & + \beta_{20} \text{CT}_i^2 \text{TCA}_i^2 + \beta_{21} \text{PH}_i^2 \text{NC}_i^2 + \beta_{22} \text{PH}_i^2 \text{TCA}_i^2 + e_i \end{aligned} \quad (7)$$

where:

Prod_i - Yield of green mass of cladodes associated with ith observation, t ha⁻¹;

- PH - plant height, cm;
- TCA - total cladode area, m²;
- NC - number of cladodes, n°;
- CA - cladode area, cm²;
- CL - cladode length, cm;
- CT - cladode thickness, cm;

CW - cladode width, cm;
 β_0 - intercept;
 $\beta_{1...n}$ - regression coefficients of the models; and,
 e_i - the error associated with the i^{th} observation.

The determination coefficient (R^2), the adjusted determination coefficient (R^2_{aj}), the Akaike Information Criterion (AIC) (Akaike, 1974), the Bayesian Information Criterion (BIC), and the selection criterion defined by the loglikelihood, which represents the logarithm value of the likelihood function considering the parameter estimates were considered for the selection of regression models.

Based on the model represented by the equation (Eq. 7), the Stepwise methodology was used to discard the less relevant variables. Regression analyzes were performed using the R software with the aid of the lm and step functions.

The regression analysis of the estimated productivity was performed with the observed values to test the predictive ability of the regression models. Subsequently, the point of intersection at the origin of the Cartesian plane was fixed, and the significance of the slope of the line was tested by the t-test, assuming as a null and alternative hypothesis the possibility of this coefficient being equal to or different from 1, respectively. Thus, if the coefficient of determination is high and the slope of the line does not differ from 1, the efficiency of prediction is assumed. The data were analyzed using the R software (R Development Core Team, 2016).

RESULTS AND DISCUSSION

The coefficients of variation and correlation values of vegetative traits with the yield of cactus pear cv. Gigante, as

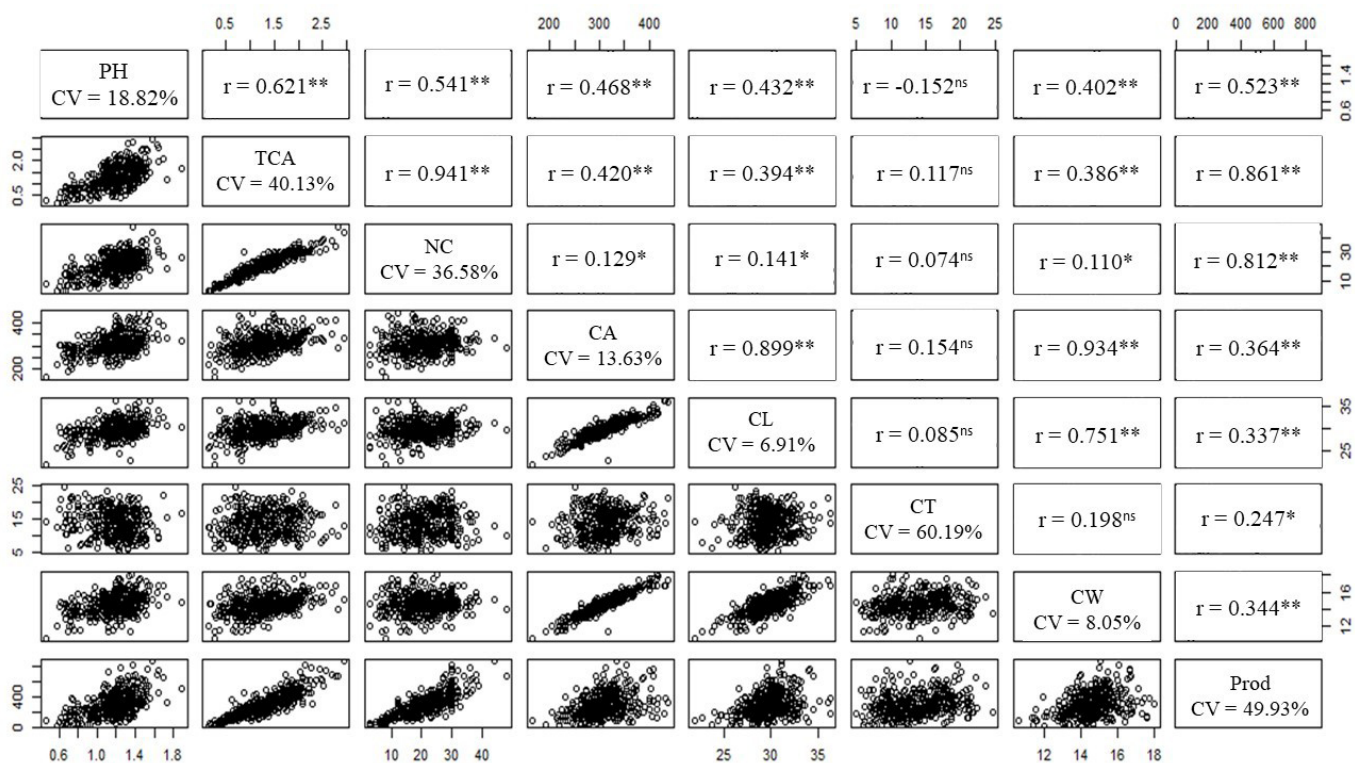
well as their significance, are shown in Figure 1. About the variability of generic traits, Gomes (2000) proposed stratifying the coefficient of variation (CV) at four categorical levels. Thus, when the range of variation is included in the classes of <10; 10.01 - 20; 20.01 - 30, and >30%, the variability is considered low, medium, high, and very high, respectively.

The CVs of the evaluated descriptors ranged between 6.91 and 60.19%, with the lowest values in the traits associated with the cladode, such as the area, length, and width of the cladode, except for the cladode thickness which showed very high variability (Gomes, 2000). Donato et al. (2014) report that the dimensions of the cladodes, especially the length and width, are determined by genotypic factors, with the low influence of the environment. However, the proper management of the crop favors the cladode thickness and, consequently, the increase in yield.

On the other hand, several studies report the wide variability of cladode thickness, as this descriptor varies over its length, although usually the thickest or central region of the cladode is measured. Also, the cladode thickness, as it is linked to photosynthetic capacity and water storage (Scalisi et al., 2016), is greatly influenced by the growth and vegetative development stage of the crop (Silva et al., 2010; Pinheiro et al., 2014; Silva et al., 2015).

The evaluated descriptors have a positive linear association with each other, which denotes, besides the high degree of relationship of vegetative variables with the yield, considerable potential of these variables to compose the prediction model.

Similarities between these results are found in Pinheiro et al. (2014) with cactus pear for all evaluated clones. Since, in this referenced study, the number of cladodes of the cactus pear expressed a high correlation with the structural traits



PH - Plant height; TCA - Total cladode area; NC - Number of cladodes; CA - Cladode area; CL - Cladode length; CT - Cladode thickness; CW - Cladode width; Prod - Yield of cactus pear cv. Gigante; ^{ns} - Not significant ($p > 0.05$); * - Significant ($0.01 < p \leq 0.05$); ** - Significant ($p \leq 0.01$)

Figure 1. Scatter plots, coefficients of variation, and the estimated correlation between vegetative traits

such as the plant height and width, with significant effects on the crop yield.

It is observed that the highest values of the correlation coefficient were associated with the traits, total cladode area, number of cladodes, and plant height, followed by variables directly related to the cladode, such as the area, length, width, and thickness (Figure 1). These results are similar to other studies on phenotypic correlation, in which, usually, the variables total cladode area and the number of cladodes express a strong relationship with the variability of cladode yield (Silva et al., 2010; 2014; Pinheiro et al., 2014; Padilha Junior et al., 2016).

By the simple linear regression procedure, compact functions were adjusted to estimate yield in cactus pear cv. Gigante, with the significance of the regression coefficients and similarity between R^2 and adjusted R^2 (Figures 2A and B).

Predictive models allow estimating yield practically and objectively in the field since it only includes an explanatory variable that is easy to determine. However, besides the best adjustment of R^2 , the equation Prod 1 has a higher predictive quality by the AIC information criterion when compared to the model Prod 2.

Models composed of variables that are easy to measure in the field are studied because they ensure practical applicability, favoring the use of the predictive tool with the insertion of values of a variable of direct measurement in the field, mainly as observed for the model of simple linear regression using the variable, number of cladodes (Figure 2). In this context, Guimarães et al. (2013) adjusted models with components of

the simple determination to estimate banana yield only by directly counting the number of hands in the bunch.

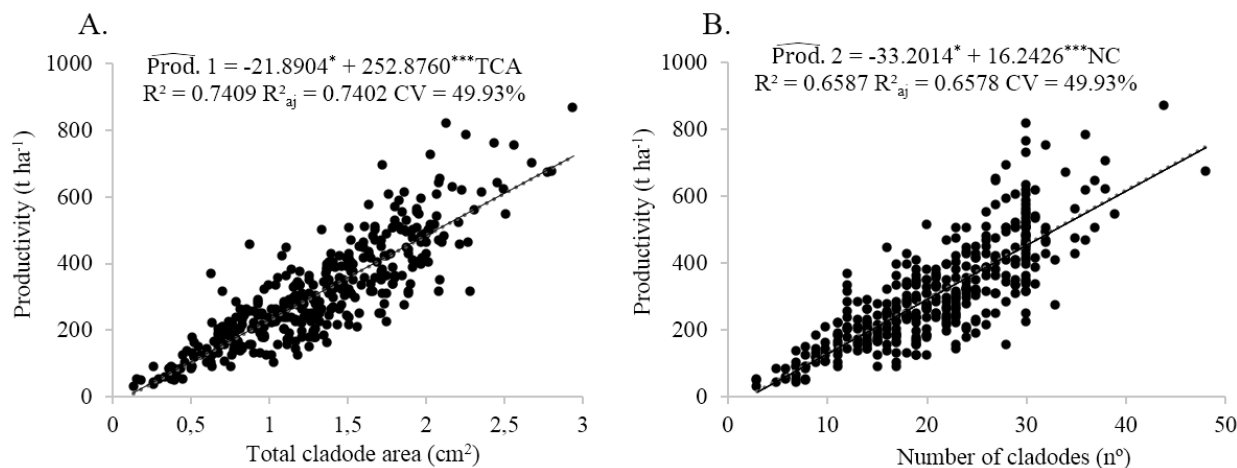
By regression analysis with the multiple linear function, models were tested to determine the yield of cactus pear cv. Gigante, according to the results presented in Table 1. The t-test for the regression models was highly significant ($p \leq 0.001$).

The yield of cactus pear cv. Gigante showed a significant correlation with all the traits analyzed (Figure 1), thus justifying the use of these variables as yield predictors. Besides the predictive capacity, the variables that make up the models presented have the advantage of direct measurement in the field in a non-destructive way (Guimarães et al., 2013; 2014).

The quality indicators AIC, BIC, and loglikelihood demonstrated that the model Prod₁ 8 has more significant potential for the prediction of cactus pear yield with R^2 of 0.7626 and $R^2_{aj} = 0.7613$ (Table 1).

However, for the tested models, the determination coefficient remained with the same approximate adjustment quality (Table 1), despite excluding vegetative traits with moderate and high correlation with yield (Figure 1), but with no significant effect to compose the model, such as plant height and the number of cladodes, respectively (Table 1). Similarly, Soares et al. (2014) and Leal et al. (2015) showed the stability of R^2 with the association of significant variables with the prediction model.

Based on the adjustment indexes of the models, presented in Table 1, and on the behavior of the equations that estimate yield, the multiple linear regression model allows to predict, in an acceptable way, the yield of the cactus pear through the



*** - Significant at $0.01 < p \leq 0.05$ and at $p \leq 0.001$ by t test, respectively

Figure 2. Estimation of productivity and the quality of adjustment by the predictive variable total cladode area (A) and the number of cladodes (B) of cactus pear cv. Gigante

Table 1. Parameters of the multiple linear regression analysis of the yield (Prod) according to the traits: PH: plant height; TCA: total cladode area; NC: number of cladodes; CA: cladode area; CL: cladode length; CT: cladode thickness; CW: cladode width

Models	Adjustment quality assessors			R^2	R^2_{aj}
	AIC	BIC	loglikelihood		
Prod 3 = $\beta_0 + \beta_1PH + \beta_2TCA + \beta_3NC + \beta_4CA + \beta_5CL + \beta_6CT + \beta_7CW + e_i$	4420.73	4456.28	-2201.36	0.7644	0.7600
Prod4 = $\beta_0 + \beta_1PH + \beta_2TCA + \beta_3CA + \beta_4CL + \beta_5CT + \beta_6CW + e_i$	4418.79	4450.39	-2201.39	0.7644	0.7606
Prod5 = $\beta_0 + \beta_1PH + \beta_2TCA + \beta_3CA + \beta_4CL + \beta_5CT + e_i$	4417.10	4444.76	-2201.55	0.7642	0.7611
Prod6 = $\beta_0 + \beta_1PH + \beta_2TCA + \beta_3CA + CT + e_i$	4415.23	4438.93	-2201.61	0.7641	0.7616
Prod7 = $\beta_0 + \beta_1PH + \beta_2TCA + \beta_3CT + e_i$	4414.31	4434.07	-2202.16	0.7634	0.7616
Prod8 = $\beta_0 + \beta_1TCA + \beta_2CT + e_i$	4413.71	4429.51	-2202.86	0.7626	0.7613

AIC - Akaike information criteria; BIC - Bayesian information criterion; R^2 - Coefficient of determination; R^2_{aj} - Adjusted determination coefficient

vegetative traits, total cladode area, and cladode thickness, with simple determination in the field (Donato et al., 2014; Padilha Junior et al., 2016), which favors the practical use of the model.

The predicted values and the observed values were listed in Figure 3A, considering the value of the slope as a determinant of the model to attest to the quality of this multiple linear regression model. This procedure is justified both by the statistical bias in search of highly significant parameters and by the need to obtain a more compact and robust model. Figure 3B represents the relationship between cactus pear yield and the predictor variables, total area of cladode, and cladode thickness. The coefficients of variation shown in Figures 3A and 3B are associated with the observed yield data. In this context, Soares et al. (2015) add that the model of the easy practical application must be composed by the smallest number of variables possible, with objective determination in the field and precise answer about the inference carried out.

Still, regarding the adjustment of the regression models expressed by the coefficient of determination (R^2), there was no difference between the equations (Table 1) regarding the predictive quality to explain the behavior of the data. However, the model Prod. 8 = $-93.7883^{**} + 247.7903^{***}TCA + 5.6736^{***}CT$ as it contains only two descriptors directly related to the cladode (TCA and CT), it becomes more simplified, adequate, and practical.

As for the indexes that define the quality of the equation adjustment, AIC, BIC, and loglikelihood, the lowest estimated values were associated with the Eq. 8 model, therefore, defined as the most appropriate (Table 1) as it presents the greatest proximity between the observed values and the estimated ones (Mello et al., 2018). Leal et al. (2015) argue about the importance of tools that measure the accuracy of the model to substantiate selection in practice.

The values of AIC, BIC, and loglikelihood are directly proportional to the sum of squares of errors. Therefore, the lower the value, the better the quality of the adjustment, defined by the smaller relative distance between the predicted and the real values (Leal et al., 2015).

Thus, the variables, being easy to determine in practice and in a direct non-destructive way, enable the researcher

or producer to estimate, with high efficiency, the yield of the cactus pear cv. Gigante. With this, it is configured as an essential tool for the success of rural planning, above all, about the size of the herd to be fed in the drought period or dry season; in which, usually, due to lack of planning, the highest mortality rate of animals occurs in the Brazilian semiarid region, compromising the economic viability of the activity and, consequently, the permanence of man in the field (Marques et al., 2017).

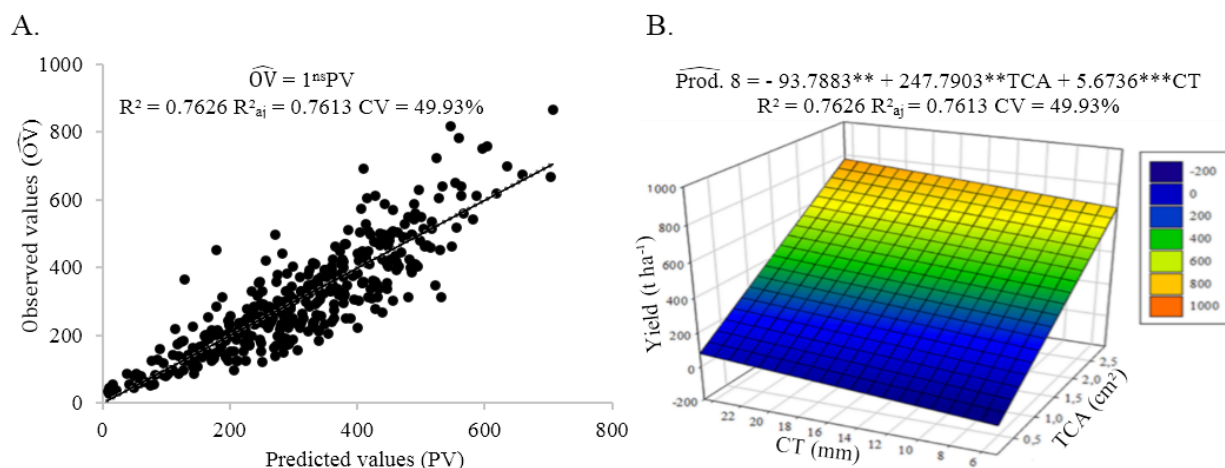
Also, it is worth considering that water is a limiting factor in animal production in regions of arid and semi-arid climates, and the use of palm in the diet of ruminants in drought periods helps animals to supply most of their water requirements (Borland et al., 2014). With this, the estimate of the productivity of the cactus pear cv. Gigante is of great importance since the possibility of predicting the food volume for ruminants achieves a dual purpose with the supply of dry matter and water.

The regression models obtained to estimate the yield of cactus pear cv. Gigante considering interactions and quadratic effects are shown in Table 2. Thus, the most appropriate model was selected according to its highest precision, which is determined by the lowest AIC value, 3290.37.

The equation Prod. 27 showed the highest predictive capacity among linear, quadratic effects and all possible interactions to estimate the yield of cactus pear cv. Gigante (Table 2), with the estimated model Prod. 27 = $9799^{**} + 9115^{*}TCA - 172.70^{*}CW - 0.127^{*}CA^2 + 3.08^{**}NC^2 - 14330^{*}TCA^2 + 376.1^{*}CLCW^2 - 2.88^{**}CLNC^2 - 12.6^{*}CWCL^2 + 0.017^{*}CWCA^2 - 1107^{**}CWTCA - 0.101^{*}CTCT^2 - 5.24^{*}PHCW^2 + 0.0105^{**}PHCA^2 - 1.83^{**}NCCW^2 + 0.0609^{*}NCCT^2 - 0.0053^{*}NCCA^2 + 0.219^{*}CL^2CW^2 - 0.065^{**}CW^2TCA^2 - 3.92^{***}CL^2TCA^2 - 0.00056^{**}CW^2CA^2 - 36.58^{**}CW^2TCA^2 + 0.0202^{***}CA^2TCA^2$.

Besides to the better suitability presented by the AIC, the R^2 and R^2_{aj} were superior to the other models with adjustments equal to 0.8187 and 0.7988, respectively, which denotes greater reliability and predictive safety (Figure 4).

Similarly, to the present study, Amaral et al. (2017) made inferences about the yield of white oats in different succession



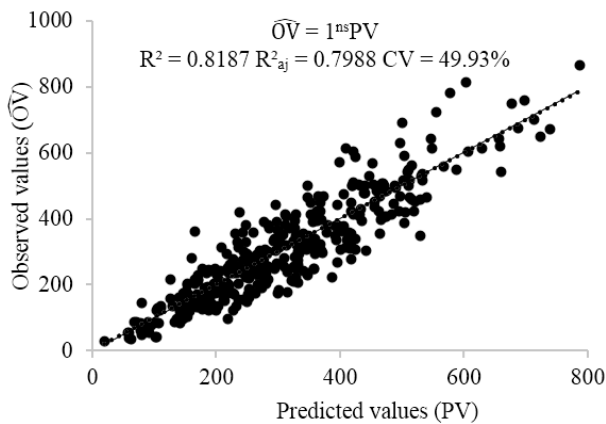
CV - Coefficients of variation associated with observed yield data; TCA - Total cladode area; CT - cladode thickness; ns - Not significant ($p > 0.05$); **, *** - Significant at $p \leq 0.01$ and at $p \leq 0.001$ by t test, respectively

Figure 3. Relationship between estimated and observed values (A) and response surface of cactus pear cv. Gigante with the vegetative traits, total cladode area, and cladode thickness (B)

Table 2. Regression model selected by the Stepwise algorithm methodology, based on the Akaike Information Criterion (AIC)

Regression models	AIC
$\text{Prod}_9 = \beta_0 + \beta_1 \text{CL}_i + \beta_2 \text{CW}_i + \beta_3 \text{CT}_i + \beta_4 \text{PH}_i + \beta_5 \text{NC}_i + \beta_6 \text{TCA}_i + \beta_7 \text{CL}_i^2 + \beta_8 \text{CW}_i^2 + \beta_9 \text{CT}_i^2 + \beta_{10} \text{CA}_i^2 + \beta_{11} \text{NC}_i^2 + \beta_{12} \text{TCA}_i^2 + \beta_{13} \text{CL}_i \text{CW}_i + \beta_{14} \text{CL}_i \text{CT}_i + \beta_{15} \text{CL}_i \text{NC}_i + \beta_{16} \text{CL}_i \text{TCA}_i + \beta_{17} \text{CL}_i \text{CL}_i^2 + \beta_{18} \text{CL}_i \text{CT}_i^2 + \beta_{19} \text{CL}_i \text{NC}_i^2 + \beta_{20} \text{CW}_i \text{CL}_i^2 + \beta_{21} \text{CW}_i \text{CA}_i^2 + \beta_{22} \text{CW}_i \text{NC}_i^2 + \beta_{23} \text{LC}_i \text{TCA}_i^2 + \beta_{24} \text{CT}_i \text{NC}_i + \beta_{25} \text{CT}_i \text{CL}_i^2 + \beta_{26} \text{CT}_i \text{CW}_i^2 + \beta_{27} \text{CT}_i \text{CT}_i^2 + \beta_{28} \text{CT}_i \text{CA}_i^2 + \beta_{29} \text{CT}_i \text{NC}_i^2 + \beta_{30} \text{PH}_i \text{NC}_i + \beta_{31} \text{PH}_i \text{CW}_i^2 + \beta_{32} \text{PH}_i \text{CA}_i^2 + \beta_{33} \text{NC}_i \text{TCA}_i + \beta_{34} \text{NC}_i \text{CW}_i^2 + \beta_{35} \text{NC}_i \text{CT}_i^2 + \beta_{36} \text{NC}_i \text{CA}_i^2 + \beta_{37} \text{NC}_i \text{NC}_i^2 + \beta_{38} \text{NC}_i \text{TCA}_i^2 + \beta_{39} \text{TCA}_i \text{CL}_i^2 + \beta_{40} \text{TCA}_i \text{CW}_i^2 + \beta_{41} \text{TCA}_i \text{NC}_i^2 + \beta_{42} \text{TCA}_i \text{TCA}_i^2 + \beta_{43} \text{CL}_i^2 \text{CW}_i + \beta_{44} \text{CL}_i^2 \text{CT}_i^2 + \beta_{45} \text{CL}_i^2 \text{CA}_i^2 + \beta_{46} \text{CL}_i^2 \text{NC}_i^2 + \beta_{47} \text{CL}_i^2 \text{TCA}_i^2 + \beta_{48} \text{CW}_i^2 \text{CT}_i^2 + \beta_{49} \text{CW}_i^2 \text{CA}_i^2 + \beta_{50} \text{CW}_i^2 \text{NC}_i^2 + \beta_{51} \text{CW}_i^2 \text{TCA}_i^2 + \beta_{52} \text{CT}_i^2 \text{CA}_i^2 + \beta_{53} \text{CA}_i^2 \text{NC}_i^2 + \beta_{54} \text{CA}_i^2 \text{TCA}_i^2 + \beta_{55} \text{NC}_i^2 \text{TCA}_i^2 + e_i$	3312.43
$\text{Prod}_{10} = \beta_0 + \beta_1 \text{CL}_i + \beta_2 \text{CW}_i + \beta_3 \text{CT}_i + \beta_4 \text{PH}_i + \beta_5 \text{NC}_i + \beta_6 \text{TCA}_i + \beta_7 \text{CL}_i^2 + \beta_8 \text{CW}_i^2 + \beta_9 \text{CT}_i^2 + \beta_{10} \text{CA}_i^2 + \beta_{11} \text{NC}_i^2 + \beta_{12} \text{TCA}_i^2 + \beta_{13} \text{CL}_i \text{CW}_i + \beta_{14} \text{CL}_i \text{CT}_i + \beta_{15} \text{CL}_i \text{NC}_i + \beta_{16} \text{CL}_i \text{TCA}_i + \beta_{17} \text{CL}_i \text{CL}_i^2 + \beta_{18} \text{CL}_i \text{NC}_i^2 + \beta_{19} \text{CW}_i \text{CL}_i^2 + \beta_{20} \text{CW}_i \text{CA}_i^2 + \beta_{21} \text{CW}_i \text{TCA}_i^2 + \beta_{22} \text{CT}_i \text{NC}_i + \beta_{23} \text{CT}_i \text{CL}_i^2 + \beta_{24} \text{CT}_i \text{CW}_i^2 + \beta_{25} \text{CT}_i \text{CT}_i^2 + \beta_{26} \text{CT}_i \text{CA}_i^2 + \beta_{27} \text{CT}_i \text{NC}_i^2 + \beta_{28} \text{PH}_i \text{NC}_i + \beta_{29} \text{PH}_i \text{CW}_i^2 + \beta_{30} \text{PH}_i \text{CA}_i^2 + \beta_{31} \text{NC}_i \text{TCA}_i + \beta_{32} \text{NC}_i \text{CW}_i^2 + \beta_{33} \text{NC}_i \text{CT}_i^2 + \beta_{34} \text{NC}_i \text{CA}_i^2 + \beta_{35} \text{NC}_i \text{NC}_i^2 + \beta_{36} \text{NC}_i \text{TCA}_i^2 + \beta_{37} \text{TCA}_i \text{CW}_i^2 + \beta_{38} \text{TCA}_i \text{NC}_i^2 + \beta_{39} \text{TCA}_i \text{TCA}_i^2 + \beta_{40} \text{CL}_i^2 \text{CW}_i^2 + \beta_{41} \text{CL}_i^2 \text{NC}_i^2 + \beta_{42} \text{CL}_i^2 \text{TCA}_i^2 + \beta_{43} \text{CW}_i^2 \text{CT}_i^2 + \beta_{44} \text{CW}_i^2 \text{CA}_i^2 + \beta_{45} \text{CW}_i^2 \text{TCA}_i^2 + \beta_{46} \text{CT}_i^2 \text{CA}_i^2 + \beta_{47} \text{CA}_i^2 \text{NC}_i^2 + \beta_{48} \text{CA}_i^2 \text{TCA}_i^2 + \beta_{49} \text{NC}_i^2 \text{TCA}_i^2 + e_i$	3303.71
$\text{Prod}_{11} = \beta_0 + \beta_1 \text{CL}_i + \beta_2 \text{CW}_i + \beta_3 \text{CT}_i + \beta_4 \text{PH}_i + \beta_5 \text{NC}_i + \beta_6 \text{TCA}_i + \beta_7 \text{CL}_i^2 + \beta_8 \text{CW}_i^2 + \beta_9 \text{CT}_i^2 + \beta_{10} \text{CA}_i^2 + \beta_{11} \text{NC}_i^2 + \beta_{12} \text{TCA}_i^2 + \beta_{13} \text{CL}_i \text{CW}_i + \beta_{14} \text{CL}_i \text{CT}_i + \beta_{15} \text{CL}_i \text{CL}_i^2 + \beta_{16} \text{CL}_i \text{NC}_i^2 + \beta_{17} \text{CW}_i \text{CL}_i^2 + \beta_{18} \text{CW}_i \text{CA}_i^2 + \beta_{19} \text{CW}_i \text{TCA}_i^2 + \beta_{20} \text{CT}_i \text{NC}_i + \beta_{21} \text{CT}_i \text{CL}_i^2 + \beta_{22} \text{CT}_i \text{CW}_i^2 + \beta_{23} \text{CT}_i \text{CT}_i^2 + \beta_{24} \text{CT}_i \text{CA}_i^2 + \beta_{25} \text{CT}_i \text{NC}_i^2 + \beta_{26} \text{CT}_i \text{CW}_i^2 + \beta_{27} \text{CT}_i \text{CT}_i^2 + \beta_{28} \text{CT}_i \text{CA}_i^2 + \beta_{29} \text{CT}_i \text{NC}_i^2 + \beta_{30} \text{PH}_i \text{CW}_i^2 + \beta_{31} \text{PH}_i \text{CA}_i^2 + \beta_{32} \text{NC}_i \text{TCA}_i + \beta_{33} \text{NC}_i \text{CW}_i^2 + \beta_{34} \text{NC}_i \text{CT}_i^2 + \beta_{35} \text{NC}_i \text{CA}_i^2 + \beta_{36} \text{NC}_i \text{TCA}_i^2 + \beta_{37} \text{TCA}_i \text{CW}_i^2 + \beta_{38} \text{CL}_i^2 \text{CW}_i^2 + \beta_{39} \text{CL}_i^2 \text{NC}_i^2 + \beta_{40} \text{CL}_i^2 \text{TCA}_i^2 + \beta_{41} \text{CW}_i^2 \text{CA}_i^2 + \beta_{42} \text{CW}_i^2 \text{TCA}_i^2 + \beta_{43} \text{CA}_i^2 \text{NC}_i^2 + \beta_{44} \text{CA}_i^2 \text{TCA}_i^2 + \beta_{45} \text{NC}_i^2 \text{TCA}_i^2 + e_i$	3292.14
$\text{Prod}_{12} = \beta_0 + \beta_1 \text{CL}_i + \beta_2 \text{CW}_i + \beta_3 \text{CT}_i + \beta_4 \text{PH}_i + \beta_5 \text{NC}_i + \beta_6 \text{TCA}_i + \beta_7 \text{CL}_i^2 + \beta_8 \text{CW}_i^2 + \beta_9 \text{CT}_i^2 + \beta_{10} \text{CA}_i^2 + \beta_{11} \text{NC}_i^2 + \beta_{12} \text{TCA}_i^2 + \beta_{13} \text{CL}_i \text{CW}_i + \beta_{14} \text{CL}_i \text{CT}_i + \beta_{15} \text{CL}_i \text{CL}_i^2 + \beta_{16} \text{CL}_i \text{NC}_i^2 + \beta_{17} \text{CW}_i \text{CL}_i^2 + \beta_{18} \text{CW}_i \text{CA}_i^2 + \beta_{19} \text{CW}_i \text{TCA}_i^2 + \beta_{20} \text{CT}_i \text{NC}_i + \beta_{21} \text{CT}_i \text{CL}_i^2 + \beta_{22} \text{CT}_i \text{CW}_i^2 + \beta_{23} \text{CT}_i \text{CT}_i^2 + \beta_{24} \text{CT}_i \text{CA}_i^2 + \beta_{25} \text{PH}_i \text{CW}_i^2 + \beta_{26} \text{PH}_i \text{CA}_i^2 + \beta_{27} \text{NC}_i \text{TCA}_i + \beta_{28} \text{NC}_i \text{CW}_i^2 + \beta_{29} \text{NC}_i \text{CT}_i^2 + \beta_{30} \text{NC}_i \text{CA}_i^2 + \beta_{31} \text{NC}_i \text{TCA}_i^2 + \beta_{32} \text{TCA}_i \text{CW}_i^2 + \beta_{33} \text{CL}_i^2 \text{CW}_i^2 + \beta_{34} \text{CL}_i^2 \text{NC}_i^2 + \beta_{35} \text{CL}_i^2 \text{TCA}_i^2 + \beta_{36} \text{CW}_i^2 \text{CA}_i^2 + \beta_{37} \text{CW}_i^2 \text{TCA}_i^2 + \beta_{38} \text{CA}_i^2 \text{NC}_i^2 + \beta_{39} \text{CA}_i^2 \text{TCA}_i^2 + e_i$	3291.18
$\text{Prod}_{27} = \beta_0 + \beta_1 \text{TCA}_i + \beta_2 \text{CW}_i + \beta_3 \text{CA}_i^2 + \beta_4 \text{NC}_i^2 + \beta_5 \text{TCA}_i^2 + \beta_6 \text{CL}_i \text{CW}_i^2 + \beta_7 \text{CL}_i \text{NC}_i^2 + \beta_8 \text{CW}_i \text{CL}_i^2 + \beta_9 \text{CW}_i \text{CA}_i^2 + \beta_{10} \text{CW}_i \text{TCA}_i^2 + \beta_{11} \text{CT}_i \text{CT}_i^2 + \beta_{12} \text{PH}_i \text{CW}_i^2 + \beta_{13} \text{PH}_i \text{CA}_i^2 + \beta_{14} \text{NC}_i \text{CW}_i^2 + \beta_{15} \text{NC}_i \text{CT}_i^2 + \beta_{16} \text{NC}_i \text{CA}_i^2 + \beta_{17} \text{CL}_i^2 \text{CW}_i^2 + \beta_{18} \text{CW}_i^2 \text{TCA}_i^2 + \beta_{19} \text{CL}_i^2 \text{TCA}_i^2 + \beta_{20} \text{CW}_i^2 \text{CA}_i^2 + \beta_{21} \text{CW}_i^2 \text{TCA}_i^2 + \beta_{22} \text{CA}_i^2 \text{TCA}_i^2 + e_i$	3290.37

(...) - equations not shown Prod. 13...26. All model coefficients were significant ($p \leq 0,01$) using the t-test



CV - Coefficients of variation associated with observed yield data; ns - Not significant ($p > 0.05$) by t-test.

Figure 4. Relationship between values estimated and observed by the model Prod. 27 with all possible interactions for the variable of the yield of cactus pear cv. Gigante.

systems with other forages. Among the models adjusted to estimate the yield of vegetable biomass and grains, the linear polynomial equations and quadratic regression reached the highest values of R^2_{aj} , above 0.87, with the highest values for quadratic models.

Although the model with quadratic interactions is composed of a higher number of predictive variables, they are easy to determine in the field to predict the yield of cactus pear cv. Gigante, which ensures the practical viability of the model, as it has been valued in studies on agricultural modeling (Guimarães et al., 2013, 2014; Soares et al., 2014; Mello et al., 2018).

CONCLUSION

The use of vegetative traits: total cladode area; the number of cladodes; area, length, thickness, and width of cladodes using multiple linear regression; quadratic interaction or only the variable, total cladode area, by using simple linear regression, allows the yield prediction of cactus pear cv. Gigante, with R^2_{aj} of 0.82, 0.76, and 0.74, respectively.

ACKNOWLEDGMENT

To the Instituto Federal do Amazonas, to the Instituto Federal Baiano, to the Universidade Estadual de Montes Claros - UNIMONTES, and to the Universidade Federal de Minas Gerais - UFMG. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001

LITERATURE CITED

Akaike, H. A new look at the statistical model identification. IEEE Transactions on Automatic Control, v.19, p.716-723, 1974. <https://doi.org/10.1109/TAC.1974.1100705>

Amaral, F. P. B.; Mantai, R. D.; Silva, J. A. G. Modelo matemático para estimativa da produtividade de biomassa e grãos de aveia. Biomatemática, v.27, p.75-88, 2017.

Bertolin, N. de O.; Filgueiras, R.; Venancio, L. P.; Mantovani, E. C. Predição da produtividade de milho irrigado com auxílio de imagens de satélite. Revista Brasileira de Agricultura Irrigada, v.11, p.1627-1638, 2017. <https://doi.org/10.7127/rbai.v11n400567>

- Borland, A. M.; Wullschlegel, S. D.; Weston, D. J.; Hartwell, J.; Tuskan, G. A.; Yang, X.; Cushman, J. C. Climate-resilient agroforestry: Physiological responses to climate change and engineering of crassulacean acid metabolism (CAM) as a mitigation strategy. *Plant, Cell & Environment*, v.38, p.1833-1849, 2014. <https://doi.org/10.1111/pce.12479>
- CODEVASF - Companhia de Desenvolvimento dos Vales do São Francisco e do Parnaíba. 2018. Dados meteorológicos. 2ª Superintendência Regional, Núcleo Avançado de Guanambi, Estação Agrometeorológicos de Ceraíma. Núcleo Avançado de Guanambi (2ª NGU). Available on: <<http://www.codevasf.gov.br/search?SearchableText=temperatura>>. Accessed on: Jan. 2018.
- Donato, P. E. R.; Pires, A. J. V.; Donato, S. L. R.; Bonomo, P.; Silva, J. A.; Aquino, A. A. Morfometria e rendimento da palma forrageira 'Gigante' sob diferentes espaçamentos e doses de adubação orgânica. *Revista Brasileira de Ciências Agrárias*, v.9, p.151-158, 2014. <https://doi.org/10.5039/agraria.v9i1a3252>
- Gomes, F. P. Curso de estatística experimental. 14.ed. Piracicaba: Degasper, 2000. 477p.
- Guimarães, B. V. C.; Donato, S. L. R.; Aspiazú, I.; Azevedo, A. M.; Carvalho, A. J. Comparison of methods for harvest prediction in 'Gigante' cactus pear. *Journal of Agricultural Science*, v.11, p.216-224, 2019. <https://doi.org/10.5539/jas.v11n14p216>
- Guimarães, B. V. C.; Donato, S. L. R.; Azevedo, A. M.; Aspiazú, I.; Silva Junior, A. A. e. Prediction of 'Gigante' cactus pear yield by morphological characters and artificial neural networks. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v.22, p.315-319, 2018. <https://doi.org/10.1590/1807-1929/agriambi.v22n5p315-319>
- Guimarães, B. V. C.; Donato, S. L. R.; Maia, V. M.; Aspiazú, I.; Coelho, E. F. Phenotypical correlations between agronomical characters in Prata type bananas (*Musa*) and its implications on yield estimate. *African Journal of Agricultural*, v.9, p.1358-1365, 2014. <https://doi.org/10.5897/AJAR12.2140>
- Guimarães, B. V. C.; Donato, S. L. R.; Maia, V. M.; Aspiazú, I.; Rodrigues, M. G. V.; Marques, P. R. R. Simple and multiple linear regressions for harvest prediction of Prata type bananas. *African Journal of Agricultural*, v.8, p.6300-6308, 2013. <https://doi.org/10.5897/AJAR2013.7544>
- Leal, A. J. F.; Miguel, E. P.; Baio, F. H. R.; Neves, D. de C.; Leal, U. A. S. Redes neurais artificiais na predição da produtividade de milho e definição de sítios de manejo diferenciado por meio de atributos do solo. *Bragantia*, v.74, p.436-444, 2015. <https://doi.org/10.1590/1678-4499.0140>
- Mantai, R.; Silva, J. A. G. da; Sausen, A. T. Z. R.; Costa, J. S. P.; Fernandes, S.; Ubessi, C. A eficiência na produção de biomassa e grãos de aveia pelo uso do nitrogênio. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v.19, p.343-349, 2015. <https://doi.org/10.1590/1807-1929/agriambi.v19n4p343-349>
- Marques, O. F. C.; Gomes, L. S. de P.; Mourthé, M. H. F.; Braz, T. G. dos S.; Pires Neto, O. de S. Palma forrageira: Cultivo e utilização na alimentação de bovinos. *Caderno de Ciências Agrárias*, v.9, p.75-93, 2017.
- Mello, M. N. de; Dias, C. T. dos S.; Martorano, L. G.; Chaves, S. S. de F.; Fernandes, P. C. C. Modelos não lineares mistos para descrever o teor de carbono orgânico no solo. *Revista Brasileira de Biometria*, v.36, p.230-240, 2018. <https://doi.org/10.28951/rbb.v36i1.176>
- Padilha Junior, M. C.; Donato, S. L. R.; Silva, J. A. da; Donato, P. E. R.; Souza, E. S. Características morfológicas e rendimento da palma forrageira 'Gigante' sob diferentes adubações e configurações de plantio. *Revista Verde de Agroecologia e Desenvolvimento Sustentável*, v.11, p.67-72, 2016. <https://doi.org/10.18378/rvads.v11i1.3710>
- Pinheiro, K. M.; Silva, T. G. F. da; Carvalho, H. F. de S.; Santos, J. E. O.; Morais, J. E. F. de; Zolnier, S.; Santos, D. C. dos. Correlações do índice de área do cladódio com características morfológicas e produtivas da palma forrageira. *Pesquisa Agropecuária Brasileira*, v.49, p.939-947, 2014. <https://doi.org/10.1590/S0100-204X2014001200004>
- R Development Core Team. The R Project for statistical computing. 2016. Available on: <<https://www.r-project.org/>>. Accessed on: Jul. 2018.
- Scalisi, A.; Morandi, B.; Inglese, P.; Bianco, R. L. Cladode growth dynamics in *Opuntia ficus-indica* under drought. *Environmental and Experimental Botany*, v.122, p.158-167, 2016. <https://doi.org/10.1016/j.envexpbot.2015.10.003>
- Soares, F. C.; Robaina, A. D.; Peiter, M. X.; Russi, J. L. Predição da produtividade da cultura do milho utilizando rede neural artificial. *Ciência Rural*, v.45, p.1987-1993, 2015. <https://doi.org/10.1590/0103-8478cr20141524>
- Soares, J. D. R.; Pasqual, M.; Lacerda, W. S.; Silva, S. O.; Donato, S. L. R. Comparison of techniques used in the prediction of yield in banana plants. *Scientia Horticulturae*, v.167, p.84-90, 2014. <https://doi.org/10.1016/j.scienta.2013.12.012>
- Silva, N. G. M.; Lira, M. de A.; Santos, M. V. F. dos; Dubeux Júnior, J. C. B.; Mello, A. C. L. de; Silva, M. da C. Relação entre características morfológicas e produtivas de clones de palma-forrageira. *Revista Brasileira de Zootecnia*, v.39, p.2389-2397, 2010. <https://doi.org/10.1590/S1516-35982010001100011>
- 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. Área do cladódio de clones de palma forrageira: Modelagem, análise e aplicabilidade. *Revista Brasileira de Ciências Agrárias*, v.9, p.633-641, 2014. <https://doi.org/10.5039/agraria.v9i4a4553>
- Silva, T. G. F. da; Primo, J. T. A.; Morais, J. E. F. de; Diniz, W. J. da S.; Souza, C. A. A. de; Silva, M. da C. Crescimento e produtividade de clones de palma forrageira no semiárido e relações com variáveis meteorológicas. *Revista Caatinga*, v.28, p.10-18, 2015.
- Taiz, L.; Zeiger, E.; Møller, I. M.; Murphy, A. Fisiologia e desenvolvimento vegetal. 6.ed. Porto Alegre: Artmed, 2017. 858p.