Australian Journal of Crop Science AJCS
AJCS 14(04):654-659 (2020) ISSN:1835-2707
doi: 10.21475/ajcs.20.14.04.p2291

Regression models for prediction of leaf area in purple ipe [*Tabebuia impetiginosa* (Mart.)]

Jéssica Sayuri Hassuda Santos*[1], Karina Tiemi Hassuda dos Santos[1], Vinicius de Souza Oliveira[2], Gleyce Pereira Santos[1], Luis Fernando Tavares de Menezes[1], Marcio Paulo Czepak[1], Antelmo Ralph Falqueto[1], Elisa Mitsuko Aoyama[1], Omar Schmildt[2], Edilson Romais Schmildt[1]

[1]Departament of Agrarian and Biological Sciences, Federal University of Espírito Santo, São Mateus-ES, Brazil
[2]Postgraduate Program in Tropical Agriculture, Federal University of Espírito Santo, São Mateus-ES, Brazil

***Corresponding author: jessicasayurihassuda@gmail.com**

Abstract

Besides its medicinal and ornamental use, *Tabebuia impetiginosa* is also very economically important. The achievement of accurate and easy-to-perform tools to determine its leaf area is fundamental for understanding the interaction between the plant and the environment. The objective of this work was to obtain regression equations by using several models that use allometric measurements of the fifth leaflet and to select the most accurate one to determine the leaf area of composite leaves of *Tabebuia impetiginosa* Mart. in a non-destructive way. By using the dimensions of the fifth leaflet such as - length (LFL in cm), maximum width (WFL in cm) and the product between LFL and WFL (LWFL) of leaf limb, the equations were estimated for linear, quadratic, potential and exponential linear models. The results showed that the determination of leaf area could be performed with excellent precision for leaves of different sizes of this species, using the product of the measurements of length and width of the fifth leaflet. The equation that best expresses the leaf area estimate of the composite leaf of *Tabebuia impetiginosa* is ELACL = 8.7772 + 2.3840 (LWFL).

Keywords: leaf area modeling, leaf dimension, non-destructive method, linear dimensions, mathematical models.
Abbreviations: CV_ coefficient of variation; ELACL_estimated leaf area of the compound leaf; LFL_length of the fifth leaflet; MAE_mean absolute error; OLACL_observed leaf area of the compound leaf; RMSE_root mean square error; WFL_ width of the fifth leaflet; length by the width of the compound leaf (LWFL).

Introduction

The species *Tabebuia impertiginosa* Mart., also known as *pau-d'arco*, purple *pau-d'arco*, *ipê-purple-de-bola*, *ipê-de-minas* belongs to the family Bignoniaceae. It is a tree-size species, reaching up to 12 m in height, with a trunk of 60-90 cm in diameter. Its leaves are compound, 5-foliolate with leathery leaflets (Lorenzi, 1992). It is native to the Amazon rainforest and tropical regions in South America and Latin America (Jin et al., 2018). Overall, the ipes are considerably economic necessary for their timber use due to the good quality of their wood, in addition to its ornamental and medicinal uses.

The leaf area determines processes related to metabolism, biomass accumulation, phenology, and crop yield due to its direct relation with photosynthesis (Demirsoy, 2009). The leaf area is fundamental for understanding the interaction between the plant and the environment (Souto et al., 2017). In order to estimate leaf area in a practical, precise, and low-cost way, leaf area is crucial mainly for non-destructive measurements over time (Oliveira et al., 2017).

Direct and indirect methods can be used to measure leaf area. For the direct method all leaves are collected, so, the method can be a destructive and labor-intensive or non-destructive, but it requires equipment that is not always accessible (Schmildt et al., 2014). The indirect and non-destructive methods, which consider previous leaf area modeling studies, allows rapid and successive evaluations in the same plant (Toebe et al., 2012), which is more advantageous in relation to the direct destructive method as more than one measurement can be made in the same leaves over time (Olfati et al., 2010).

Modeling studies determine regression equations (Demirsoy, 2009), using allometric measurements of the leaves (Olfati et al., 2010). This type of study has been done for several forest species, among them *Fagus sylvatica* L. (Anev et al., 2016), *Alnus subcordata*, *Populus deltoides* and *Taxodium distichum* (Eslamdoust et al., 2017), *Carpotroche brasiliensis* (Brito- Rocha et al., 2016), *Helicteres isora* L. and *Vitex negundo* L.(Kumbhani et al., 2017), *Ulmus japonica* (Cai et al., 2017), *Valeriana jatamansi* Jones (Walia and Kumar, 2017) and *Erythroxylum simonis* Plowman (Ribeiro et al., 2018). However, in *Tabebuia impertiginosa*, the same one utilized in this work we can find the the study carried out by Monteiro et al. (2017) who worked on the modeling of leaf area of seedlings, using the average width of leaflets.

Nevertheless, no studies could be found in the literature for *ipê* that used only one leaflet, thus the objective of this work was to obtain equations in linear and non-linear models using the length and width of only one leaflet, the fifth, to estimate the leaf area of the entire leaf of adult plants of *Tabebuia impertiginosa*.

Results and Discussion

Leaf allometric measurements

The descriptive statistics of the allometric measurements of purple-ipe leaves are shown in Table 1. It can be seen that the total amplitude, standard deviation, and coefficient of variation of all the characteristics in the leaves used to obtain the equations was higher than those of the validation. As a result, all values of these leaves are between the maximum and minimum of those used to obtain the equations, with no extrapolation. Levine et al. (2017) recommend that no extrapolation of the independent variable should occur when using regression equations to estimate values.

The coefficient of variation (CV) of all the allometric measurements of the leaves to estimate the equations was considered high by the classification of Pimentel-Gomes (2009), therefore, reinforcing that the leaves used in this work were of different sizes. That was also observed by Ribeiro et al. (2018), when modeling the leaf area in *Erythroxylum simonis* Plowman.

Mathematical models

Figure 1 shows the OLACL dispersion diagrams as a function of the allometric measurements LFL, WFL, and LWFL. It was observed through the distribution of the data trend of simple linear adjustment for OLACL as a function of LWFL based on the dispersion of the values, that there is not a good fit for OLACL as a function of LFL and for OLACL as a function of WFL in the simple linear model. However, this visual finding needs to be investigated by statistical criteria (Levine et al., 2017). Hence, the estimation of the equations based on the linear simple, quadratic, potential, and exponential mathematical models (Table 2) will be presented in a sequence.

The observations made by the visualization of the dispersion diagrams are confirmed by the estimated equations (Table 2). Most of the equations obtained in the experiment were well adjusted to R^2 higher than 0.90, except for only E3. It is observed that the best fit using only one measure of the fifth leaflet, either length (LFL) or width (WFL), was the potential model, showing values of coefficient of determination (R^2) higher than 0.97. On the other hand, the largest values of R^2 were 0.9906 of the equations L3 and Q3, which were simple and quadratic linear type, respectively, using the product of length by width (LWFL).

Criteria for choosing the best equation

Regarding the selection of an equation, one should not only consider R^2 as a measure of precision in modeling but rather by interpreting statistical measures in validation (Walia and Kumar, 2017). Thus, by observing the criteria of the linear coefficients (β_0) statistically equal to zero and the angular coefficient (β_1) statistically equal to one, only the equations L3 and Q3 are considered adequate for estimation of leaf area of the whole leaf using measurements of the fifth leaflet. Both equations use LWFL as an allometric measure (Table 3). Santana et al. (2018) also found in the quadratic model a suitable adjustment of the leaf area of the compound leaf from the central leaflet in six legume species. The need for using the product of both measures of the leaflet to estimate the leaf area of the composite leaf was also observed for other species with compounds leaves such as soybean (Richter et al., 2014), common bean (Lakitan et al., 2017) and pigeon pea (Pezzini et al., 2018).

By taking into account the ease of use, the simple linear equation is recommended: ELACL = 8.7772 + 2.3840(LWFL), which has in the other validation criteria, values of R^2, the Willmott index (d), MAE and RMSE very close to those obtained by the quadratic equation (Q3). The best fit recommended for estimating the leaf area of the compound leaf using a simple linear model based on the product between leaf length and leaflet length measurements was also observed by other authors (Richter et al., 2014; Lakitan et al., 2017).

The adjustment of the recommended above mentioned simple linear equation, as well as the visualization of the validation criteria for the adjusted equation, are shown in Figure 2.

Monteiro et al. (2017) obtained a potential equation to estimate the leaf area of *T. impetiginosa*, in leaves of seedlings, where the average width of the leaflets is required. By using a simple linear equation where the measurement of the width and length of only one leaflet is required, in this case, the fifth one, facilitates and streamlines the analysis, thus reducing workforce and time of evaluations.

Materials and methods

Plant materials

For the *Tabebuia impertiginosa* leaf area experiment, 136 leaves were collected from 10 adult plants at eight years of age located in the Campus of the Federal University of Espírito Santo (UFES), in São Mateus-ES (latitude of 18º40'36 ''S, longitude 39º51'35''W), where compound leaves of all sizes were removed.

The selected leaves were those that did not present any damages or attacks of diseases by pests. They were randomly selected according to the method of Benincasa (2004) who states that when the area where a more significant number of plants can be randomly harvested, the number of plants to be harvested should be between 10 and 20 plants as values below 10 may induce errors and above 20 do not significantly increase sampling accuracy. The leaves were then taken to the Plant Improvement Laboratory of UFES for the analysis. Of the 163 leaves collected, 110 were used to obtain the equations and 53 for their validation.

Mathematical models

In order to obtain the allometric measurements, the purple-ipe leaves were scanned with the aid of an HP Deskjet F2050® scanner and estimated using ImageJ® version 1.32j

Table 1. Number of composite leaves (n), mean, minimum, maximum, total amplitude (TA), standard deviation (SD) and coefficient of variation (CV) for length (LFL), width (WFL), length by width (LWFL) of the fifth leaflet and observed leaf area of the composite leaf (OLACL) of adult plants of purple-ipe, used for estimation of regression equations and validation.

Statistic	LFL (cm)	WFL (cm)	LWFL (cm^2)	OLACL (cm^2)
Estimation of regression equations				
n	110	110	110	110
Mean	14.02	6.20	104.55	258.01
Minimum	2.10	1.10	2.31	9.42
Maximum	28.50	13.30	379.05	960.60
TA	26.40	12.20	376.74	951.18
SD	6.18	2.95	88.47	211.90
CV (%)	44.10	47.57	84.62	82.13
Validation				
n	53	53	53	53
Mean	14.52	6.76	106.15	260.57
Minimum	5.90	2.60	15.34	41.44
Maximum	21.00	12.90	265.74	615.24
TA	15.10	10.30	250.40	573.38
SD	3.80	2.38	60.61	145.14
CV (%)	26.18	35.26	57.09	56.09

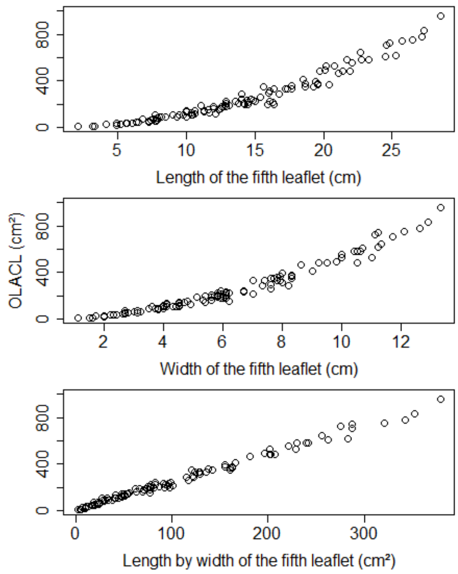

Fig 1. Dispersion diagram for relation between the measurements of the limb of the fifth leaflet and the observed leaf area in the composite leaf (OLACL).

Table 2. Regression models for the estimation of leaf area of the composit leaf (ELACL, cm²) from measurement of the fifth leaflet (x) of *Tabebuia impetiginosa*, and respective coefficients of determination (R^2).

Model[1]	Equation	R^2 (%)
L1	ELACL = -202.4433 + 32.8385(LFL)	0.9180
L2	ELACL = -173.8113 + 69.6797(WFL)	0.9399
L3	ELACL = 8.7772 + 2.3840(LWFL)	0.9906
Q1	$\mathrm{ELACL} = 17.4028 - 2.4027(\mathrm{LWF}) + 1.1698(\mathrm{LWFL})^2$	0.9735
Q2	$\mathrm{ELACL} = -0.0018 + 6.9562(\mathrm{LWF}) + 4.5705(\mathrm{LWFL})^2$	0.9815
Q3	$\mathrm{ELACL} = 9.8043 + 2.3607(\mathrm{LWF}) + 0.00008(\mathrm{LWFL})^2$	0.9906
P1	$\mathrm{ELACL} = 1.3782(\mathrm{LFL})^{1.9154}$	0.9723
P2	$\mathrm{ELACL} = 7.8094(\mathrm{WFL})^{1.8313}$	0.9811
P3	$\mathrm{ELACL} = 3.2187(\mathrm{LWFL})^{0.9453}$	0.9901
E1	$\mathrm{ELACL} = 20.7020\,(1.1642)^{\mathrm{LFL}}$	0.9104
E2	$\mathrm{ELACL} = 24.5606(1.3722)^{\mathrm{WFL}}$	0.9081
E3	$\mathrm{ELACL} = 63.2413(1.0098)^{\mathrm{LWFL}}$	0.7691

[1] L1, L2 and L3 = linear simples; Q1, Q2 and Q3 = quadratic; P1, P2 and P3 = power; E1, E2 and E3 = exponential. The index 1, 2 and 3 indicate the length, width and length x width of the fifth leaflet, respectively.

Table 3. Linear coefficient ($\hat{\beta}_0$), slope ($\hat{\beta}_1$), coefficient of determination (R^2, %), mean absolute error (MAE), root mean square error (RMSE) and index Willmott (WILLMOTT, 1981) obtained from regression adjusted between the estimated leaf area of composite leaf (ELACL, dependent variable) and the observed leaf area of sample of 53 leaves composite leaf (OLACL, independent variable) of *Tabebuia impetiginosa.*

Model(1)	$\hat{\beta}_0$ (2)	$\hat{\beta}_1$ (3)	R^2	MAE	RMSE	d
L1	67.2679^{**}	0.7949^{**}	0.8660	44.1235	55.8648	0.9554
L2	7.7893^{ns}	1.1109^{**}	0.9555	41.1250	52.9678	0.9710
L3	6.1903^{ns}	0.9811^{ns}	0.9848	13.1803	17.8758	0.9961
Q1	45.1734^{**}	0.7697^{**}	0.8980	35.1851	52.3564	0.9593
Q2	-33.2444^{*}	1.2074^{**}	0.9324	37.2652	59.5734	0.9662
Q3	6.3004^{ns}	0.9794^{ns}	0.9848	13.2787	17.9103	0.9961
P1	51.7892^{**}	0.7431^{**}	0.8959	36.2453	54.3741	0.9547
P2	-33.7467^{*}	1.2133^{**}	0.9364	37.0557	59.3621	0.9666
P3	10.7936^{*}	0.9664^{*}	0.9851	13.3963	18.0069	0.9960
E1	15.8434^{ns}	0.7832^{**}	0.8913	47.9763	64.8407	0.9402
E2	-160.1456^{**}	1.7129^{**}	0.7860	78.9703	167.4795	0.8370
E3	-48.2866^{**}	1.0094^{ns}	0.8638	61.4824	73.9535	0.9399

(1) L1, L2 and L3 = linear simples; Q1, Q2 and Q3 = quadratic; P1, P2 and P3 = power; E1, E2 and E3 = exponential. The index 1, 2 and 3 indicate the length, width and length x width of the fifth leaflet, respectively. (2) *, ** Linear coefficient different to zero by t-test at 5% and 1% probability. ns Non-significant. (3) *, ** Slope different to one by t-test at 5% and 1% probability. ns Non-significant.

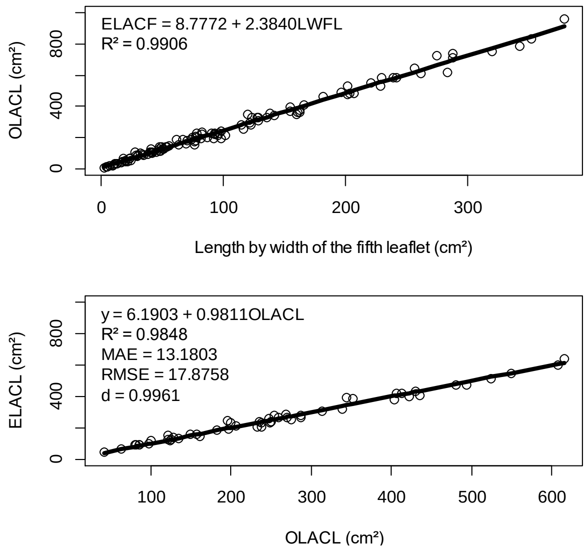

Fig 2. Representation of leaf area of the composite leaf in the *Tabebuia impetiginosa* Mart. by simple linear model, for the product between length and width of the fifth leaflet: A = equation estimated in 110 leaves; B = model validation with 53 leaves.

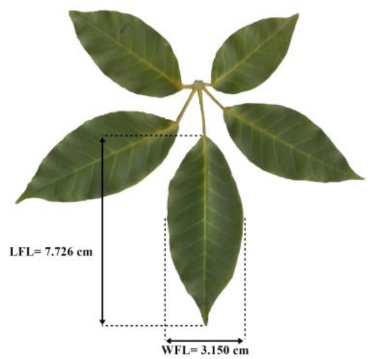

Fig 3. Illustration of a composite leaf of purple- ipe emphasizing the fifth leaflet from which the length and width were measured for modeling the leaf area of the composite leaf.

software (Wayne Rasband National Institute of Health, USA). As shown in Figure 3, the measurements were performed on the central leaflet of the compound leaf (fifth leaflet) in the longest direction (LFL, in cm) on the main vein and largest width (WFL, cm) made perpendicularly as well as the observed leaf area of the compound leaf (OLACL, in cm^2).

Validation

The OLACL dispersion diagram was plotted as a function of the allometric measurements of the fifth leaflet for the sample of 110 compound leaves. Equations of linear and non-linear equations were estimated (simple and quadratic) and parameters, respectively. OLACL was used as the dependent variable (axis y), as a function of length (LFL), width (WFL) or product of length-by-width (LWFL), which were independent variables (x-axis). The estimated leaf area models of the compound leaf (ELACL) are linear, quadratic, potential and exponential simple linear type, represented by equations 1, 2, 3 and 4, respectively.

(1) $\mathrm{ELACL} = \hat{\beta}_0 + \hat{\beta}_1 x$
(2) $\mathrm{ELACL} = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2$
(3) $\mathrm{ELACL} = \hat{\beta}_0 x^{\hat{\beta}_1}$
(4) $\mathrm{ELACL} = \hat{\beta}_0 \hat{\beta}_1^{x}$

From the three allometric measurements of the leaves, three equations were obtained for each model, totaling 12 equations to estimate the leaf area of the compound leaf. The parameters β_0 and β_1 were estimated through the method of the least squares. Besides, the linearization of the potential and exponential functions were previously carried out.

The validation of the 12 compound leaf area estimation equations (ELACL) was performed based on the estimated values of the sample of 53 compound leaves. For each equation, a simple linear regression ($y = \hat{\beta}_0 + \hat{\beta}_1 \mathrm{OLACL}$) of the relationship between ELACL and OLACL was adjusted, and the coefficient of determination was obtained. For the adjustment of the referred simple linear regression, the least squares method was used. The following hypothesis was tested: $H_0: \beta_0 = 0$ versus $H_0: \beta_0 \neq 0$ versus $H_0: \beta_1 \neq 1$, employing Student's t-test at 5% probability of error. Then, the mean absolute error (MAE) (equation 5), the root mean square error (RMSE) (equation 6), and the Willmott index (Willmott, 1981) (equation 7) were determined.

(5) $MAE = \frac{\sum_{i=1}^{n}(ELACL_i - OLACL_i)^2}{n}$

(6) $RMSE = \sqrt{\frac{\sum_{i=1}^{n}(ELACL_i - OLACL_i)^2}{n}}$

(7) $d = 1 - \left[\frac{\sum_{i=1}^{n}(ELACL_i - OLACL_i)^2}{\sum_{i=1}^{n}(|ELACL_i - \overline{OLACL}| + |OLACL_i - \overline{OLACL}|)^2}\right]$

$\mathrm{ELACL_i}$ is the estimated values of the leaf area of the compound leaves i; $\mathrm{OLACL_i}$ is the observed values of leaf area of the i compound leaves; $\overline{\mathrm{OLACL}}$ is the average of the observed values of the compound leaves; n is the sample size for validation, where n = 53, in this work.

The best equation to estimate compound leaf area as a function of LFL, WFL and LWFL was selected through objective validation criteria based on the linear coefficient (β_0) not statistically different from zero, angular coefficient (β_1) not statistically different from one, MAE and RMSE closer to zero, R^2 and Willmot index (d) closer to one. The best-fit equation, combined with its validation, was presented graphically. Statistical analysis and graphical representation were performed using R software (R Core Team, 2018).

Conclusions

The most adequate linear equation to estimate the compound leaf area of *Tabebuia impertiginosa* Mart., is ELACL = 8.7772 + 2.3840 (LWFL), where LWFL is the product of length by the width of the fifth leaflet of the compound leaf.

Acknowledgments

We thank the National Council for Scientific and Technological Development (CNPq-processes 311181/2018-0) and the Coordination for the Improvement of Higher Education Personnel (Capes) for granting scholarships. To the Laboratory of Restinga and Atlantic Forest Ecology by technical support.

References

Anev S, Dimitrova-Mateva P, Lamlom S, Chaneva G, Tzvetkova N (2016) Non-destructive allometric method for estimation of leaf area in common beech (*Fagus sylvatica* L.). Forestry Ideas. 22(2): 198–205.

Benincasa MMP (2004) Análise de crescimento de plantas: noções básicas. Jaboticabal: FUNEP. 42.

Brito-Rocha E, Anjos L, Schilling AC, Dalmolin AC, Mielke MS (2016) Individual leaf area estimations of a dioecious tropical tree species *Carpotroche brasiliensis* (Raddi) A. Gray, Achariaceae. Agrofor Syst. 91: 9-15.

Cai H, Di X, Jin G (2017) Allometric models for leaf area and leaf mass predictions across different growing seasons of elm tree (*Ulmus japonica*). J Forest Res. 28(5): 975-982.

Demirsoy H (2009) Leaf area estimation in some species of fruit tree by using models as non-destructive method. Fruits. 64: 45-51.

Eslamdoust J, Sohrabi H, Hosseini SM, Naseri B (2017) Leaf biomass and leaf area equations for three planted trees in Iran. Eur J Biol Res. 7(1): 50-58.

Jin Y, Jeong KM, Lee J, Zhao J, Choi S, Baek K (2018) Development and validation of an analytical method readily applicable for quality control of *Tabebuia impetiginosa* (Taheebo) ethanolic extract. J AOAC Int. 101(3): 695-700.

Kumbhani NR, Kuvad RP, Thaker VS (2017) Development of linear model for leaf area measurement of two medicinally important plants: *Helicteres isora* L. and *Vitex negundo* L. J App Biol Biotech. 5(03): 57-60.

Lakitan B, Widuri LI, Meihana M (2017) Simplifying procedure for a non-destructive, inexpensive, yet accurate trifoliate leaf area estimation in snap bean (*Phaseolus vulgaris*). J Appl Hortic. 19(1): 15-21.

Levine DM, Stephan DF, Szabat KA (2017) Statistics for managers using Microsoft® Excel: global edition (8th ed.), London: Pearson.

Lorenzi H (1992) Árvores brasileiras: manual de identificação e cultivo de plantas arbóreas nativas do Brasil. Nova Odessa: Ed. Plantarium. 352.

Monteiro EB, Silva AC, Souza AP, Silva CC, Kazama VS, Tanaka AA (2017) Statistical parameters to estimate the leaf area of native forest seedlings of genus *Tabebuia* and *Handroanthus.* Biosci J. 33(4): 956-967.

Olfati JA, Peyvast GH, Shabani H, Nosratie-rad Z (2010) An estimation of individual leaf area in cabbage and broccoli using nondestructive methods. J Agri Sci Technol. 12: 627-632.

Oliveira PS, Silva W, Costa AAM, Schmildt ER, Vitória EL (2017) Leaf area estimation in litchi by means of allometric relationships. Rev Bras Frutic. 39 (spe): 1-6.

Pezzini RV, Cargnelutti Filho A, Alves BM, Follmann DN, Kleinpaul JA, Wartha CA, Silveira DL (2018) Models for leaf area estimation in dwarf pigeon pea by leaf dimensions. Bragantia. 77(2): 221-229.

Pimentel-Gomes F (2009) Curso de estatística experimental (15th. ed.). Piracicaba, SP: Fealq.

R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ribeiro, JES, Barbosa, AJS, Albuquerque, MB (2018) Leaf Area Estimate of *Erythroxylum simonis* Plowman by Linear Dimensions. Floram. 25(2): e20170108.

Richter GL, Zanon AJ, Streck NA, Guedes JVC, Kräulich B, Rocha TSM, Winck JEM, Cera JC (2014) Estimating leaf area of modern soybean cultivars by a non-destructive method. Bragantia. 73(4): 416-425.

Santana HA, Rezende BR, Santos WV, Silva AR (2018) Models for prediction of individual leaf area of forage legumes. Rev Ceres. 65(2): 204-209.

Schmildt ER, Hueso JJ, Cuevas J (2014) Allometric models for determining leaf area of vine 'Sugraone'. Cienc Tec Vitivinic. 29(10): 61-81.

Souto AGL, Cordeiro MHM, Rosado LDS, Santos CEM, Bruckner CH (2017) Non-destructive estimation of leaf area in passion fruit (*Passiflora edulis* L.). Aust J Crop Sci. 11(12):1534-1538.

Toebe M, Cargnelutti Filho A, Burin C, Fick AL, Neu IMM, Casarotto G, Alves, BM (2012) Modelos para a estimação da área foliar de feijão de porco por dimensões foliares. Bragantia. 71(1): 37-41.

Walia S, Kumar R (2017) Development of the Nondestructive Leaf Area Estimation Model for *Valeriana* (*Valeriana jatamansi* Jones). Commun. Soil Sci Plan. 48(1): 83-91.

Willmott CJ (1981) On the validation of models. Phys Geogr. 2(2): 184-194.