Abstract

Studies of leaf surface are important for the ecophysiological context. We tested 12 allometric equations to estimate the leaf area of the endemic species *Erythroxylum pauferrense* Plowman, using different morphometric measurements (n = 600). The analysis of variance (p < 0.001), determination coefficient (R^2_adj), standard error of estimates (S_{xy}), and graphical interpretation of the dispersion of errors were used to select the models. The residual bias was estimated through the confidence interval using t-Student distribution (p = 0.01). Simple linear equations (Ŷ = 0.0244 + 0.7204 × CxL) and power functions (Ŷ = 0.7279 × CxL^{0.9971}) were validated from an independent sample of leaves, confirming their accuracy. The results are important to guide actions of conservation of the endemic species.

Keywords: Erythroxylaceae, allometry, leaf area.

1. INTRODUCTION AND OBJECTIVES

Leaf area is a key characteristic of ecophysiological investigations, since the study of leaves is essential for understanding the functioning of ecosystems by the consistency of their chemical, structural and physiological properties (Díaz et al., 2016; Wright et al., 2004). Besides, leaves are a fundamental unit in processes such as light interception, temperature regulation, and water balance (Ehleringer, 2017), and its measurement is important to evaluate plant development and their photosynthetic processes (Liu et al., 2017). Temperature, climatic and microclimatic factors may limit the anatomy and architecture of leaves (Wright et al., 2004). In ecological studies, the leaf area is used to evaluate the effects of management, competition among species, and to conservation, providing information about the adaptation of morphological and functional groups of plants in the most diverse habitats (Cornelissen et al., 2003; Dainese et al., 2015; Pérez-Harguindeguy et al., 2013; Sfair et al., 2016).

The methods for estimation of leaf area are classified as destructive and non-destructive. Models based on leaf area estimation using linear leaf dimensions of several species are common (Bianco et al., 2004; Coelho Filho et al., 2005; Maracajá et al., 2008; Schmildt et al., 2016; Silva et al., 2013) and promising in relation to more expensive methods of measurement; besides, it is a reliable alternative for the evaluation of the area of many leaf species. However, a good estimation by the models depends on sample range (all leaf size variation), morphological characteristics (e.g., the shape of the lamina, number of lobes, leaflets), and should represent the complete allometric expansion pattern of a species. Modeling to obtain leaf area estimation emerges as an excellent tool for forest science field.

Information on the ecophysiology of species such as *Erythroxylum pauferrense* Plowman are scarce. *E. pauferrense* is an endemic species of Paraíba, Northeast region of Brazil; it can reach 1.5 m to 4 m in height, inhabits the understory vegetation and develops under limited light conditions (Loiola et al., 2007; Plowman, 1986). The species was recently discovered in botanical (Loiola et al., 2007) and phytosociological terms (Andrade et al., 2006; Barbosa et al., 2004; Queiroz et al., 2006), and the ecophysiological aspects are little known. The taxon has an endemic distribution restricted to one of the devastated sectors of the Atlantic Forest – of the Alto-Montana forest – and is listed as a critically endangered species partially due to its restricted phytogeographic distribution.
The subpopulations of *E. pauferrense* remain isolated and the number of individuals is declining due to habitat degradation (CNCFlora, 2012). Studies aiming at monitoring and conserving the genetic resources of the species should help its recovery and conservation. Therefore, ecophysiological studies, such as leaf area-related, indicate the species’ performance against its habitat and subsidizes conservation actions for endemic genetic resources.

We expect that information at the organism level will add value to the knowledge of the species ecology and contribute to the conservation of native forest heritage, as well as fostering new researches. This work aimed to select and validate the equation that best estimates and fits the data of leaf area for *E. pauferrense*. Therefore, we tested different allometric models using several leave morphometric measures and selected the best fit model according to the accuracy, precision, and reliability.

### 2. MATERIALS AND METHODS

#### 2.1. Characterization of the area

The study was performed in Mata do Pau-Ferro State Park (coordinates: 6°58’12”S, 35°42’15”W), located in the city of Areia-PB, Brazil (Figure 1). The park has approximately 600 ha, is the largest remnant of *Brejo de Altitude* in the state of Paraíba, thus, it is an essential habitat for the preservation of *E. pauferrense* (Barbosa et al., 2004). The phytosociology is a type of Ombrophilous Dense Forest, that is, Open Ombrophilous Forest (IBGE, 2012).

The relief of the region is characterized as wavy (slopes ≤8% to 20%) to strongly wavy (slopes of 20% to 45%). The climate is as according to Köppen-Geiger classification, described as hot and humid with autumn and winter rains (Kottek et al., 2006).

#### 2.2. Data sampling

Thirty individuals of *E. pauferrense* were selected throughout the study area and 600 leaves were collected (healthy and free from any damage caused by herbivory), covering all the intraspecific variation of leaves in the area (from regenerating individuals to adult). The different canopy levels of the understory were also used as a selection criterion for leaves. Sampling was made from August 2015 to September 2016.

The leaves were stored in dark, moistened plastic bags (Pérez-Harguindeguy et al., 2013) and analyzed at the Laboratory of Plant Ecology (LEV) at the Federal University of Paraiba, Campus II (UFPB). All leaves were listed and then scanned through a digital scanner, and the images were processed through Image-Pro Plus® software version 4.0 (Media Cybernetics). Leaf area was measured in square millimeter (mm²) and converted to square centimeters (cm²). The maximum length (*L*\(_{\text{max}}\)) was measured as the greatest distance between the point of insertion of the petiole in the leaf limbus and the apex of the leaf. The maximum width (*W*\(_{\text{max}}\)) was measured as the largest horizontal dimension to the main vein axis (Figure 2).

![Figure 2. Scheme of the measures of leaf lamina of *E. pauferrense*. *L*\(_{\text{max}}\): maximum length; *W*\(_{\text{max}}\): maximum width.](image)

---

**Figure 1.** Map of the location of the study area, State Park of Mata do Pau-Ferro, Areia, PB.
2.3. Theoretical models

Twelve models were tested, using combinations between the dimensional variables $L_{\text{max}}$, $W_{\text{max}}$ and the leaf area values. All models were elaborated by the principle of simplicity and practicality. The leaf limbus was taken as the leaf area, the dependent variable in the models. The variables $L_{\text{max}}$ and $W_{\text{max}}$ and the product of the linear combination “$L_{\text{max}} \times W_{\text{max}}$” of each leaf were used as independent variables (Table 1).

2.4. Statistical analysis

The equations were fitted in the software R version 3.4.0 (“R Development Core Team”, 2019). Model selection was performed based on the criteria: (i) significance of the analysis of variance ($F$ test; $p \leq 0.01$); (ii) the value of the coefficient of determination adjusted by the degrees of freedom ($R^2_{\text{adj}}$); (iii) the standard error of the estimate in absolute terms ($S_{\text{xy}}$); (iv) the mean square of residuals (MSresid.); and (v) the graphical pattern of errors dispersion in percentage terms ($\epsilon_i$%). The dispersion pattern of the residues was verified for the total sample data following the statistical procedures for graphic interpretation (Graybill, 2000; Zuur et al., 2009).

For each model, the average deviation was tested (if equal to or different from zero, null and alternative hypotheses, respectively) (Equation 13) to infer the model accuracy. Errors close or equal to zero mean better model accuracy. Then, the confidence interval (CI) for the mean was calculated using the t-Student distribution ($p > 0.01$).

$$\left( \frac{\sum_{i=1}^{600} c_i}{n} = 0 \text{ and } \frac{\sum_{i=1}^{600} c_i}{n} \neq 0 \right)$$

The combination of described criteria indicates the best-fitted equation for the estimation of leaf area.

2.5. Model validation

The selected models were submitted to a validation process. For the validation, a second independent sample of leaves was randomly collected ($n = 200$) during the same period of the main sample ($n = 600$). The leaf area, length and width of the new sample were measured according to the methodology previously used. Subsequently, the estimated and observed values of leaf area were compared using a Student’s t-test at 99% of confidence level ($p = 0.01$). The model validation with the independent sample of leaves aims to prove the stability and the significance of the parameters (Zar, 2010).

3. RESULTS

3.1. Leaf characteristics

The leaf area ranged from 0.13 cm² to 19.24 cm², with an average of 7.75 cm² and standard deviation of 4.51 cm². A strong correlation between the linear measurements of the leaf and the observed leaf area was found (i.e., $R^2_{\text{adj}}$ from 0.9095 to 0.9977; $p < 0.001$) (Table 2). The equations 3, 6 and 12 best fitted the data showing the lower values of standard error of the estimation ($S_{\text{xy}}$) (Table 2).

| Table 1. Theoretical models to predict the leaf area of the E. pauferrense species and their respective generalizations. |
| --- |
| Equation | Type | Allometric models |
| (1) | Linear | $Y = \beta_0 + \beta_1 \times L_{\text{max}} + e_i$ |
| (2) | Linear | $Y = \beta_0 + \beta_1 \times W_{\text{max}} + e_i$ |
| (3) | Linear | $Y = \beta_0 + \beta_1 \times (L \times W) + e_i$ |
| (4) | Power | $Y = \beta_0 \times L_{\text{max}}^{\beta_1} + e_i$ |
| (5) | Power | $Y = \beta_0 \times W_{\text{max}}^{\beta_1} + e_i$ |
| (6) | Power | $Y = \beta_0 \times (L \times W)^{\beta_1} + e_i$ |
| (7) | Modified Power | $Y = L_{\text{max}}^{\beta_1} + e_i$ |
| (8) | Modified power | $Y = W_{\text{max}}^{\beta_1} + e_i$ |
| (9) | Modified power | $Y = (L \times W)^{\beta_1} + e_i$ |
| (10) | Modified linear | $Y = \beta_1 \times L_{\text{max}} + e_i$ |
| (11) | Modified linear | $Y = \beta_1 \times W_{\text{max}} + e_i$ |
| (12) | Modified linear | $Y = \beta_1 \times (L \times W) + e_i$ |

Y: leaf area; $L_{\text{max}}$: maximum length; $W_{\text{max}}$: maximum width; LW: product $L_{\text{max}} \times W_{\text{max}}$; $\beta_0$ and $\beta_1$: model parameters; $e_i$: random error.

| Table 2. Adjusted models for leaves of (E. pauferrense) with the respective regression coefficients ($\hat{\beta}_0$, $\hat{\beta}_1$), standard errors of estimates ($S_{\text{xy}}$), adjusted R-squared ($R^2_{\text{adj}}$), the degree of freedom of residuals (DFresid), mean square of residuals (MSresid) and the dependent variable, leaf area ($\hat{Y}$) as a function of maximum length ($L_{\text{max}}$), maximum width ($W_{\text{max}}$) and the product $L_{\text{max}} \times W_{\text{max}}$ (L x W). |
| --- |
| Model | Coefficients | $S_{\text{xy}}$ | $R^2_{\text{adj}}$ | DFresid | MSresid | $P$ | Estimator=\hat{Y} |
| --- |
| (1) | $-3.2606$ | $2.7114$ | $1.1179$ | $1.1179$ | $0.9387$ | $598$ | $1.2497$ | $<0.001$ | $\hat{Y} = 3.2606 + 2.7114 L_{\text{max}}$ |
| (2) | $-4.4021$ | $5.2192$ | $1.3479$ | $0.9109$ | $598$ | $1.8168$ | $<0.001$ | $\hat{Y} = 4.4021 + 5.2192 W_{\text{max}}$ |
| (3) | $0.0244$ | $0.7204$ | $0.2239$ | $0.9975$ | $598$ | $0.0502$ | $<0.001$ | $\hat{Y} = 0.0244 + 0.7204 (L \times W)$ |
| (4) | $0.6477$ | $1.7058$ | $0.8393$ | $0.9654$ | $598$ | $0.7044$ | $<0.001$ | $\hat{Y} = 0.6477 L_{\text{max}}^{1.7058}$ |
The modified models 10, 11 and 12 showed high coefficient of determination (0.96, 0.94 and 0.99, respectively) (Table 2), but the error differed from zero (Figure 3) by both the hypothesis test and the confidence interval, showing a bias in the distribution. The residuals of the modified power equations (Equations 7, 8 and 9) also differed from zero, therefore rejecting the null hypothesis of absence of bias (Figure 3). The linear models 1 and 2, power models 4 and 5, modified power models 7, 8 and 9, and the modified linear models 10 and 11 were discarded due to the bias of the residuals. A more detailed analysis of the results of the equations indicated a bias from 18.33% to 99.33%, considering errors ≥40%. The leaf area was underestimated by the equations 9 and 12 and overestimated by the equations 4, 5, 7, 8, 10 and 11, therefore these models were rejected. Besides, the residuals showed a non-homogeneous distribution (heteroscedasticity).

The equations 3 (linear model for relation with the product $L_{\text{max}} \times W_{\text{max}}$) and 6 (power model for relationship relation with the product $L_{\text{max}} \times W_{\text{max}}$) best fitted the data according to the analysis of deviation (Figure 3). The graphical interpretation of the residue dispersion analysis was important to verify the differences between the statistical parameters of equations (Zuur et al., 2009). A difference of estimation between the equations 3 and 6 was not found (Figure 4). Both equations showed significant determination coefficients ($R^2_{\text{adj}}$) and low standard errors of estimates ($S_{xy}$) (Table 2), besides a low relative dispersion of residuals (Figure 5). The relative residues range from −26.29% to 11.22% (Figure 5a and b), are normal and constant, and reflect the unexplained variability for the measures of the leaves of *E. pauferrense* ($= 7.5 \pm 4.51 \text{ cm}$, mean and standard deviation, respectively). Equations using the length and width product ($L \times W$) to describe leaf area are common for species with simple leaves (Cardozo et al., 2014; Maracajá et al., 2008).

The low dispersion of the values of leaf area suggests a close correlation with the product $L \times W$, justifying the good adjustment of the equations 3 and 6 (Figure 5a and b). The brackets highlight adjustment details and accuracy of the data around the mean for both large and small leaves.

Using the validation analysis, it was verified that the estimated values of leaf area were close to the observed for both linear (3) and power models (6), thus, a significant statistical difference between estimated and observed values was not found (Table 3). The models 3 and 6 are the most suitable for the estimation of leaf area of *E. pauferrense* for all size class, from initial to fully developed leaves.

**Table 2.** Continued...

| Model | Coefficients | $S_{xy}$ | $R^2_{\text{adj}}$ | $DF_{\text{resid}}$ | $MS_{\text{resid}}$ | P | Estimator = $\hat{Y}$ |
|-------|--------------|---------|--------------------|---------------------|-------------------|---|---------------------|
| (5)   | 1.1907       | 2.0653  | 0.9594             | 0.9095              | 598               | <0.001 | $\hat{Y} = 1.1907 W_{\text{max}}^{2.0653}$ |
| (6)   | 0.7279       | 0.9971  | 0.9775             | 0.9975              | 598               | <0.001 | $\hat{Y} = 0.7279 (L \times W)^{0.9971}$ |
| (7)   | -            | 1.4459  | 0.9537             | 0.9421              | 599               | <0.001 | $\hat{Y} = 1.4459 L_{\text{max}}^{1.4459}$ |
| (8)   | -            | 2.2233  | 0.9569             | 0.8773              | 599               | <0.001 | $\hat{Y} = 2.2233 W_{\text{max}}^{2.2233}$ |
| (9)   | -            | 0.882   | 0.9282             | 0.991               | 599               | <0.001 | $\hat{Y} = 0.8820 (L \times W)^{0.8820}$ |
| (10)  | -            | 2.0178  | 0.9649             | 0.9649              | 599               | <0.001 | $\hat{Y} = 2.0178 L_{\text{max}}^{2.0178}$ |
| (11)  | -            | 3.5396  | 0.9410             | 0.9761              | 599               | <0.001 | $\hat{Y} = 3.5396 W_{\text{max}}^{3.5396}$ |
| (12)  | -            | 0.7221  | 0.2241             | 0.9977              | 599               | <0.001 | $\hat{Y} = 0.722 (L \times W)$ |

**Figure 3.** Asymmetry of the deviations for each allometric equation (1 to 12). The lower limit of the box indicates the 25th percentile; the lines inside the boxes denote the median; the upper limit of the box indicates the 75th percentile. The vertical lines above and below the box indicate the minimum and maximum values (outliers are shown as points). Asterisks (*) denote biased equations, according to confidence intervals calculated for the mean of the deviations (99% by Student’s t-distribution).
4. DISCUSSION

The regression equations that estimate better the leaf area of *Erythroxylum pauferrense* were the simple linear (3) and power (6) models using the product of the maximum length and maximum width (L × W), which showed normality in the error distribution (homoscedasticity) and good adjustment for all statistical parameters. Our results agree with other studies using plants with simple leaves (Lima et al., 2012; Santos, 2016; Toebé et al., 2010). The product between maximum length and maximum width has proven to be a valid measure of leaf area also for the species *Vernonia ferruginea* (Less.) (Souza & Amaral, 2015), *Psidium cattleianum* and *Psychotria suterella* (Moraes et al., 2013); but some studies found lower coefficient of determination for the relation between leaf and L × W for the species *Crotalaria juncea* L. (Cardozo et al., 2014). The results show the efficiency of the model with the product of the length and width as the predictor variable for the determination of leaf area, which may be related to a higher amount of information provided through the two dimensions (Boeger et al., 2009; Monteiro et al., 2005).

The equations determined by the simple linear models (1 and 2, using the variables maximum length and maximum width, respectively) were rejected despite their significance,
also reported as inadequate by other studies since the relation with these variables does not expand in a completely linearly way (Williams & Martinson, 2003). All single dimension models, those that use only the length or width of leaves, have high values for the $R^2$ adj, but they were rejected due to the high residue deviation (Figure 3). Some researchers suggest that the use of the single-dimension models (Antunes et al., 2008; Pompei et al., 2012) which may reduce the amount of time and resources implemented in experiments despite the simplification. However, the use of single dimension models is not recommended to estimate leaf area of *E. pauferrense*.

Monteiro et al. (2005) found good estimations using the length and width of the leaves carried out separately for cotton leaf (*Gossypium hirsutum* L.). However, similar to our results, Schmildt et al. (2017) verified a better fit using the product between length and width for the leaf area of cacao tree (*Theobroma cacao* L.). Once again, the mathematical models using $L \times W$ are effective by presenting higher values of $R^2$. These results corroborate some studies with other species of different leaf morphological profiles (Cardozo et al., 2014; Maracajá et al., 2008) and are more evident in simple leaves with elliptic and oblong limbs. Likewise, the power function in our study gives accurate and reliable estimates leaf area for the species *E. pauferrense*.

In this work, it was chosen to dispense with any mathematical transformation of the variables for the simple and practical use of the equations. This criterion and the acceptance of the normality of the residue offer credibility for the selected models. Leaves removal is only necessary to estimate the allometric models. Once the allometric equation is known, the leaf area is determined using a graduated ruler or a measuring tape with no need of leaves removal.

The processes of development and growth of forest species are important in the ecological context, since the structural condition of an individual is related to the environmental and genetic factors that contributed to its growth from the seedling stage (Silva et al., 2015).

Therefore, studies using allometric models for the estimation of leaf area provide a tool for data measure for evolutionary and adaptation processes (Brito-Rocha et al., 2016; Forrester et al., 2017; Zhang & Pan, 2011). Our study, thus, produces data that may contribute to the preservation of the natural genetic resources of the species *E. pauferrense* in its habitat facing the biological risk scenario of the Atlantic Forest.

5. CONCLUSIONS

The equations indicated to estimate leaf area of *Erythroxylum pauferrense* are the linear model ($\hat{Y} = 0.0244 + 0.7204 (L \times W)$) and the power function ($\hat{Y} = 0.7279 (L \times W)^{0.9971}$) using the product of maximum length and maximum width as the predictor variable. These equations show great precision, the absence of bias and high reliability. The indirect measure of leaf area through the product of maximum length and maximum width provides a simple non-destructive method of estimation and avoids the use of expensive equipment.

The estimation of leaf area using the product of maximum length and maximum width of leaves is also a consensus for commercial species. Therefore, we expect that this study provides a tool to the simplification of various biological research (e.g., experimental, ecological and ecophysiological data) and to understand aspects of biological responses of the species to natural ecosystems or conservation actions.

ACKNOWLEDGEMENTS

To the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), for the scholarship.

SUBMISSION STATUS

Received: 8 Nov. 2017
Accepted: 10 June 2018

CORRESPONDENCE TO

Ana Jéssica Soares Barbosa
Universidade Federal da Paraíba (UFPB), Programa de Pós-Graduação em Agronomia, Centro de Ciências Agrárias, Campus II, Rodovia BR 079, km 12, CEP 58397-000, Areia, PB, Brasil.
e-mail: ajsoaresbarbosa22@gmail.com

FINANCIAL SUPPORT

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

REFERENCES

Almeida SMZ, Soares AM, Castro EM, Vieira CV, Gajego EB. Alterações morfológicas e alocação de biomassa em plantas jovens de espécies florestais sob diferentes condições de sombreamento. Ciência Rural 2005; 35(1): 62-68. 10.1590/S0103-84782005000100010

Andrade LA, Oliveira FX, Nascimento IS, Fabricante JR, Sampaio EVSB, Barbosa MRV. Análise florística e estrutural de matas ciliares ocorrentes em brejo de altitude no município de Areia, Paraíba. Revista Brasileira de Ciências Agrárias 2006; 1(1): 31–40.

Antunes WC, Pompei MFD, Carretero FMD. Allometric models for non-destructive leaf area estimation in coffee (*Coffea arabica* and *Coffea canephora*). Annals of Applied Biology 2008; 153(1): 33–40. 10.1111/j.1744-7348.2008.00235.x

Barbosa MRV, Agra, MF, Sampaio EVSB, Cunha JP, Andrade LA. Diversidade florística da Mata do Pau-Ferro, Areia, Paraíba. In: Porto KC, Cabral JFP, Tabarelli M, editors. *Brejos de altitude em Pernambuco e Paraíba: história natural, ecologia e conservação*. Brasilia: Ministério do Meio Ambiente; 2004. p. 111-122. (Série Biodiversidade 9).
Silva GH, Queiroz JE, Neto AGS. Avaliação da área foliar de três espécies florestais ocorrentes no semiárido paraibano (Amburana cearensis, Caesalpinia ferrea, Caesalpinia pyramidalis). Revista de Biologia e Farmácia 2013; 9(3): 1-11.

Silva RAR, Vieira FA, Fajardo CG, Araújo FS. Padrões alométricos da palmeira carnaúba (Copernicia prunifera (Mill.) H. E. Moore). Nativa 2015; 3(1): 56-58. 10.14583/2318-7670.v03n01a09

Souza MC, Amaral CL. Non-destructive linear model for leaf area estimation in Vernonia ferruginea Less. Brazilian Journal of Biology 2015; 75(1): 152-156. 10.1590/1519-6984.09813

Ehleringer J. Fotossíntese: considerações fisiológicas e ecológicas. In: Taiz L, Zeiguer E, Møller I, Murphy A, editors. Fisiologia e desenvolvimento vegetal. 6th ed. Porto Alegre: Artmed; 2017. p. 245-268.

Toebe M, Brum B, Lopes SJ, Carnealutti Filho A, Silveira TR. Estimativa da área foliar de Crambe abyssinica por discos foliares e por fotos digitais. Revista Ciência Rural 2010; 40(2): 475-478. 10.1590/S0103-84782010000200036

Williams L, Martinson TE. Nondestructive leaf area estimation of 'Niagara' and 'DeChaunac' grapevines. Scientia Horticulturae 2003; 98(4): 493-498. 10.1016/S0304-4238(03)00020-7

Wright IJ, Reich PB, Westoby M, Ackerly DD, Baruch Z, Bongers F et al. The worldwide leaf economics spectrum. Nature 2004; 428(6985): 821-827. 10.1038/nature02403

Zar JH. Bioestatistical analysis. 5th ed. New Jersey: Pearson; 2010.

Zhang L, Pan L. Allometric models for leaf area estimation across different leaf-age groups of evergreen broadleaved trees in a subtropical forest. Photosynthetica 2011; 49(2): 219-226. 10.1007/s11099-011-0027-x

Zuur AF, Ieno EN, Elphick CS. A protocol for data exploration to avoid common statistical problems. Methods in Ecology & Evolution 2009; 1(1): 3-14. 10.1111/j.2041-210X.2009.00001.x