Allometric models to estimate peanuts leaflets area by non-destructive method

João Everthon da Silva Ribeiro[1,*] (iD) , Ester dos Santos Coêlho[1] (iD) , Pablo Henrique de Almeida Oliveira[1] (iD) , Welder de Araújo Rangel Lopes[1] (iD) , Elania Freire da Silva[1] (iD) , Anna Kézia Soares de Oliveira[1] (iD) , Lindomar Maria da Silveira[1] (iD) , Daniel Valadão Silva[1] (iD) , Aurélio Paes Barros Júnior[1] (iD) , Thiago Jardelino Dias[2] (iD)

1. Universidade Federal Rural do Semi-Árido ROR – Mossoró (RN), Brazil.
2. Universidade Federal da Paraíba ROR – Areia (PB), Brazil.

ABSTRACT: The determination of leaf area is fundamental for studies related to plant growth and physiology. Thus, non-destructive methods allow an accurate estimate of the leaf area through linear dimensions of the leaves. The research objective was to construct allometric equations to estimate the leaflet area of peanut cultivars. Then, 2,605 leaflets were collected from six peanut cultivars (IAC Caiapó, IAC 8112, Runner IAC 886, BRS Havana, BRS 151 L7, and IAC Tatuí), with more than 400 leaflets sampled for each cultivar. We measured the length, width, product between length and width, and leaflet area. Linear and non-linear models (linear, linear without intercept, power, and exponential) were built, and the best equation was chosen using the statistical criteria: highest coefficient of determination (R^2), Pearson's linear correlation coefficient (r), Willmott's agreement index (d), lowest Akaike information criterion (AIC), and root mean square of the error (RMSE). It was found that the models that used the product between length and width were the most suitable for estimating the leaflet area of peanut cultivars. Given the little intraspecific morphological variability, it was possible to group the cultivars, and model $\hat{y} = 0.875 * LW^{0.929}$ was indicated to estimate the peanut leaflet area accurately, regardless of the cultivar.

Key words: *Arachis hypogaea* L., biometry, leaf morphotypes, nonlinear models.

INTRODUCTION

The nutritional importance and productive potential of peanut (*Arachis hypogaea* L.) classifies this crop as a multipurpose oilseed, offering several medicinal and economic benefits (Akram et al. 2018). The chemical and nutritional composition of peanuts includes vitamins, folic acid, thiamine, and tocopherols (Akram et al. 2018). In addition, its seeds are sources of fatty acids, with great importance for the composition of its oil (Toomer 2017). There is a great diversity of peanut genotypes, which differ in oil quality and morphological structure (Krishna et al. 2015).

Because of the relevance of the peanut crop, studies related to growth and development are of great need to estimate the productive potential of its different genotypes (Akram et al. 2018). Among the aspects required to indicate plants' growth status and physiological behavior, leaf area is characterized as an essential parameter (Keramatlou et al. 2015).

The leaf area directly influences the vital processes of plants, which include gas exchange and the interception of irradiance (Córcoles et al. 2015). The shape and amount of intercepted light is a determining factor for photosynthesis and directly interferes with acquiring essential resources for the formation of carbohydrates (Mattos et al. 2020). Therefore, the light energy captured by the leaves of a crop is critical to shaping its growth (Liu et al. 2021).

For physiological and ecophysiological studies, the measurement of leaf area provides an estimate of the transpiration rate and net absorption and ensures an understanding of the interaction between the growth of the species and its developmental environment (Sabouri and Sajadi 2022). Carrying out studies with leaf area modeling involves processes of plant community composition, evolution, competition, and adaptation and provides clarification of fruit quality attributes (Keramatlou et al. 2015).

Methods for estimating leaf area can be direct or indirect and differ in plants destructive or non-destructive sampling (Zhang 2020). Direct methods are economically unfeasible, limited by logistical factors, and prevent successive measurements of leaves (Hernandéz-Fernandez et al. 2021). On the other hand, a non-destructive indirect method of determining leaf area is through allometric relationships and considers the proportionality of the linear dimensions of the leaves (Santos et al. 2021).

To monitor leaf development and quantify a regression model that estimates leaf area, it is necessary to use digital processing methods, viable, accurate, and economic tools for crops (Sauceda-Acosta et al. 2017). Using this approach, Sabouri and Sajadi (2022) evaluated the use of image processing and regression methods to estimate the leaf area of chia (*Salvia hispanica* L.), quinoa (*Chenopodium quinoa* Willd.), and bitter melon (*Momordica charantia* L.).

In addition, the approach to leaf area estimation by these methods covers studies with pepper (Yeshitila and Taye 2016), cactus forage (Lucena et al. 2021), cocoa (Salazar et al. 2018), fava beans (Peksen 2007), and mango (Ghoreishi et al. 2012). However, for a critical evaluation of models to estimate the leaf area, it is necessary to use many leaves or leaflets to validate the model (Hernandéz-Fernandez et al. 2021).

For peanuts, there is a lack of studies that contemplate a more significant number of cultivars and demonstrate morphological differences and provide a regression model that facilitates subsequent studies with this species. We formulated that the linear and non-linear models are reliable for estimating the leaflet area (LA) of peanut cultivars. Thus, the objective of this work was to propose an equation that accurately estimates the leaf area of peanut cultivars through linear dimensions of the leaflets using several linear and non-linear models.

MATERIAL AND METHODS

Plant material and experimental conditions

The experiment was conducted in an experimental area of the didactic garden belonging to the Agricultural Sciences Center of the Universidade Federal Rural do Semi-Árido, RN, Brazil (5°12'25.26"S, 37°19'6.42"W). The region's climate is classified as BSh, being dry and very hot, with a dry and rainy season (Alvares et al. 2013). It has an average temperature of 28 °C and annual rainfall of around 695 mm. The soil in the region is classified as Eutrophic Red-Yellow Argisol (Embrapa 2018).

For the construction of the models, peanut cultivars with an erect growth habit were planted, with sowing in September 2021 and harvesting in December 2021. Each experimental plot consisted of a cultivar sown in a 15-m planting row, with the density of 15 seeds per linear meter and spacing of 0.90 m between rows.

Plant sampling and image processing

At 85 days after planting, a total of 2,605 leaflets (n) mature, expanded, free from damage or pests, were collected from six peanut cultivars [1: IAC Caiapó (n = 427); 2: IAC 8112 (n = 433); 3: Runner IAC 886 (n = 422); 4: BRS Havana (n = 441); 5: BRS 151 L7 (n = 465); and 6: IAC Tatuí (n = 417)] (Fig. 1), with more than 400 leaflets randomly sampled for each cultivar. We selected leaflets in different shapes and sizes to test the model's generality and seek greater variability in the data. The leaves were placed in plastic envelopes and kept in the shade immediately after collection to maintain turgidity.
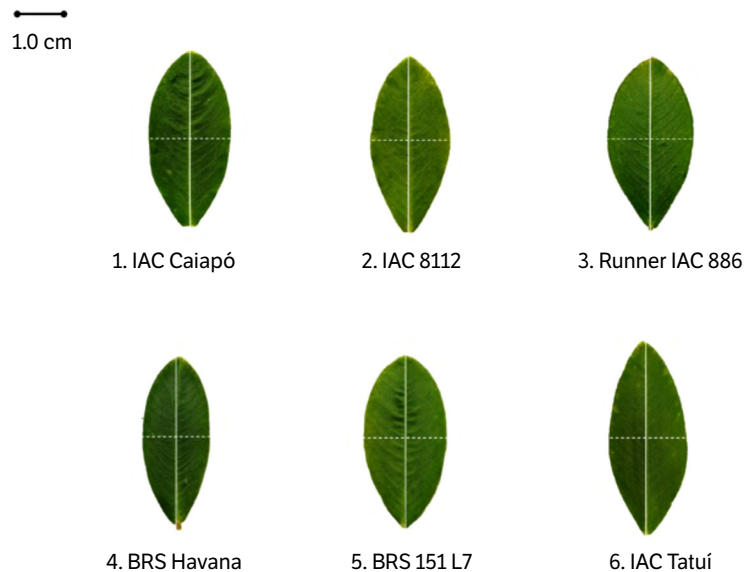
Figure 1. Representative leaflets of peanut cultivars. Sampling information is shown in Fig. 2. The vertical line represents the length on each leaf, and the dashed line represents the leaflet width. The scale bar refers to 1 cm.

The leaflets were separated from the leaves and then digitalized in a scanner (Epson model L395, Tokyo, Japan) with a resolution of 1,200 × 1,200 dpi, with the images processed and analyzed individually using the ImageJ software (National Institutes of Health, United States of America). The digitalized images included three rulers as reference indicators for measurements. We calculated the maximum length (L) (distance between the petiole insertion end and the opposite distance from the midrib), and maximum width (W) (largest measure perpendicular to the midrib), and LA were calculated (cm^2) for each leaflet. Then, we calculated the product between length and width (LW) (cm^2).

Data analysis

To compare the foliar parameters between the cultivars (L, W, LW, and LA), a unidirectional analysis of variance was performed, and then the honestly significant difference (HSD) was performed using the Tukey's test at the level of 5% probability (Zar 1996).

To verify whether there is accuracy in the estimates of the regression coefficients between L and W, we evaluated the degree of collinearity between these parameters. Therefore, we calculated the variance inflation factor (VIF) (Eq. 1) (Marquardt 1970) and the tolerance value (T) (Eq. 2) (Gill 1986). In this case, if the VIF is greater than 10 and the T is less than 0.1, it indicates that L and W have multicollinearity to affect the estimate of the leaflet area and, consequently, should be disregarded one of these two leaf parameters (L or W) for fitting regression models in the prediction of leaflet area (Gill 1986).

$$VIF = \frac{1}{1 - r^2} \tag{1}$$

$$T = \frac{1}{VIF} \tag{2}$$

in which: r: Pearson's linear correlation coefficient between L and W.

To estimate leaflet area based on linear dimensions, tests were performed with 90 linear and non-linear regression models between LA (dependent variable) and length, width, and product (LW) (independent variables) of each cultivar, and later they were tested with the grouped data (all the grouped cultivars). Among the models tested, 25 presented

satisfactory criteria for estimating the leaflet area. Logarithmic and polynomial models from second to fifth order were excluded. Thus, for greater data precision and speed in the analysis, 10 equations from four models were used for each cultivar and the grouped data.

The regression models used were linear ($\hat{y} = \beta_0 + \beta_1 * x + \varepsilon_i$), linear without intercept ($\hat{y} = \beta_1 * x + \varepsilon_i$), power ($\hat{y} = \beta_0 * x^{\beta_1} + \varepsilon_i$), and exponential ($\hat{y} = \beta_0 * \beta_1^x + \varepsilon_i$). Therefore, the value of yˆ estimated the LA as a function of x, which corresponds to the linear dimensions of the leaflets (L, W, and LW). To validate the models, we randomly collected 350 independent leaflets among the cultivars during the same leaflet collection period, to estimate the leaflet area.

To analyze whether the model was chosen to accurately estimate the leaflet area for all cultivars with distinct leaf characteristics, the hypotheses (H_0) $\beta_0 = 0$ *versus* (H_a) $\beta_0 \neq 0$ and (H_0) $\beta_1 = 0$ *versus* (H_a) $\beta_1 \neq 1$ were tested. From Student's t-test at 5% probability, β_0 and β_1 are regression coefficients.

The criteria for choosing the best model to estimate the peanut leaflet area were: higher coefficient of determination (R^2) (Eq. 3), more elevated Pearson's linear correlation coefficient (r) (Eq. 4), lower Akaike information criterion (AIC) (Eq. 5), higher Willmott's agreement index (d) (Eq. 6), and lower root mean square error (RMSE) (Eq. 7).

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y'_i)^2} \tag{3}$$

$$r = \frac{\sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2 \ \sum_{i=1}^{n} (x_i - \bar{x})^2}} \tag{4}$$

$$AIC = -2\ln L(x\backslash\hat{\theta}) + 2(p) \tag{5}$$

$$d = 1 - \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n} (|\hat{y}'_i| + |y'_i|)^2} \tag{6}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}i - yi)^2}{n}} \tag{7}$$

in which: $\hat{y}_i$: estimated values of LA; y_i: observed values of LA; $\bar{y}_i$: average of observed values; $\hat{y}'_i = \hat{y}_i - \bar{y}$; $y'_i = y_i - \bar{y}$; $L(x\backslash\hat{\theta})$: maximum likelihood function; p: number of model parameters; n: observation numbers; x_i and y_i: ith observations of the variables y and x; $\bar{y}$ and $\bar{x}$: means of variables y and x.

Descriptive analyses were performed to calculate maximum and minimum values, median, and coefficient of variation (CV), and data normality was verified by the Shapiro-Wilk's test. The observed and estimated LA were compared using Student's t-test for paired samples (p < 0.01). We performed statistical data analysis with the software R® v.4.1.2 (R Core Team 2022).

RESULTS

The high number of leaflets of the cultivars used in this study (n = 2,605) provided a high data variation for constructing models to estimate the peanut LA (Fig. 2). The leaflets of the cultivars showed a length range of 0.82-8.38 cm (IAC Caiapó), 1.99-8.59 cm (IAC 8112), 0.98-8.50 cm (Runner IAC 886), 1.20-8.45 cm (BRS Havana), 1.44-8.64 cm (BRS 151 L7), and 1.78-8.58 cm (IAC Tatuí) (Fig. 2). The LW ranges were 0.60-4.34, 0.68-3.86, 0.63-4.60, 0.67-3.81, 0.96-4.11 and 0.91-4.07 cm for the same cultivars (Fig. 2). Regarding the LA of the cultivars, averages of 12.2 ± 0.26, 11 ± 0.21, 10.6 ± 0.24, 11.8 ± 0.21, 11.7 ± 0.27, and 11.1 ± 0.22 cm2 were observed, with coefficients of variation above 30% (Fig. 2).
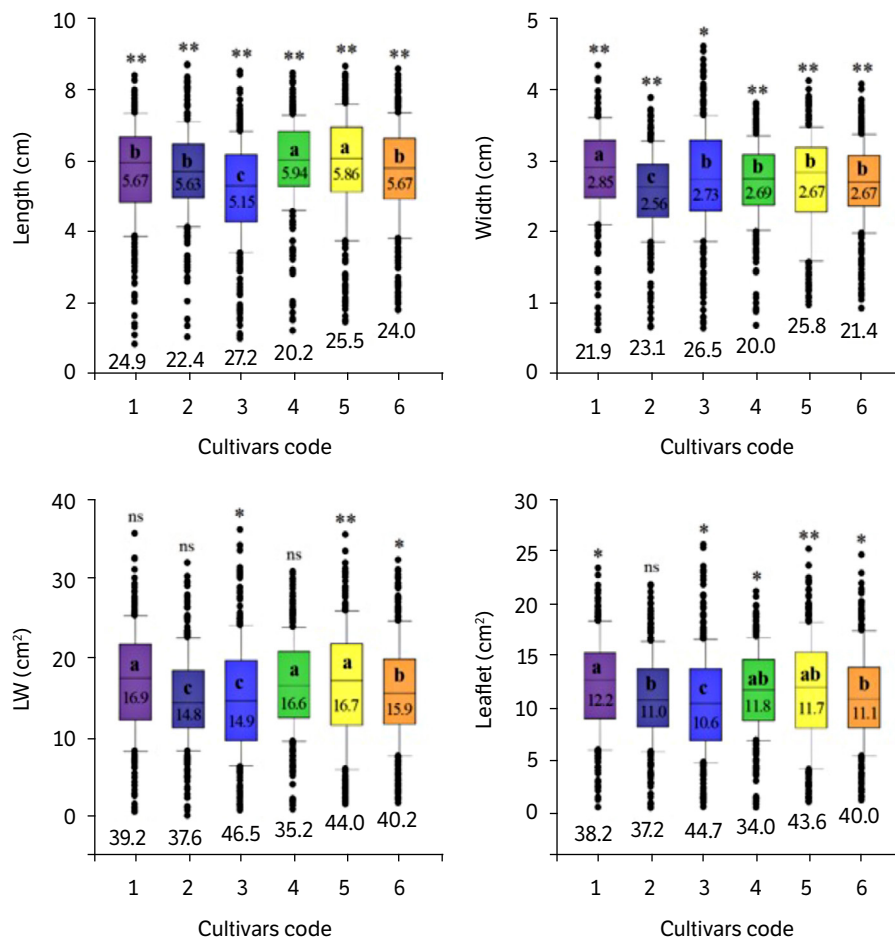
*Significant at 5% probability; **significant at 1% probability; nsnot significant; 1: IAC Caiapó; 2: IAC 8112; 3: Runner IAC 886; 4: BRS Havana; 5: BRS 151 L7; 6: IAC Tatui.
Figure 2. Descriptive analysis of (a) length, (b) width, (c) product between length and width, and (d) leaflet area for peanut cultivars. The upper and lower ends of each box represent the ¼ and ¾ quartiles; the horizontal line of each box refers to the median. The distribution of upper and lower points and the bars represent extreme values (maximum and minimum) of the data set. The numbers inside the boxes represent the averages of the cultivars. The numbers below the dots refer to the coefficients of variation. In each panel, cultivars with the same letter do not differ from each other by Tukey's test at the level of 5% probability. The significance above the points refers to the Shapiro-Wilk's normality test.

Regarding the comparison of L, W, LW, and LA between the cultivars, significant differences were observed in all leaf parameters, such as the cultivar IAC Caiapó, which had greater W, LW, and LA, while its L was the second largest, and the cultivar Runner IAC 886 had the most petite L and size of leaflets compared to the other cultivars (Fig. 2). The data for L and W did not adjust to the normal distribution in all cultivars, while for LW the data showed normal distribution in the cultivars IAC Caiapó, AIC 8112, and Runner IAC 886. In LA, there was an adjustment of the normal distribution only for the cultivar IAC 8112 (Fig. 2).

Linear and non-linear association patterns between L, W, LW, and LA were observed in the data used to build the regression model to estimate the peanut LA (Fig. 3). We observed linear patterns between LW and LA, non-linear patterns between L and LA, and between W and LA, thus evidencing the need to use different models for data adjustment and validation (Fig. 3). According to the preliminary analysis for model calibration, the VIF values ranged from 1.66 to 8.24, and the T values ranged from 0.12 to 0.60, respectively. Thus, for all cultivars, VIF values were < 10 and T was > 0.10, showing that the collinearity between L and W is considered negligible, and we can include these parameters in the regression models.
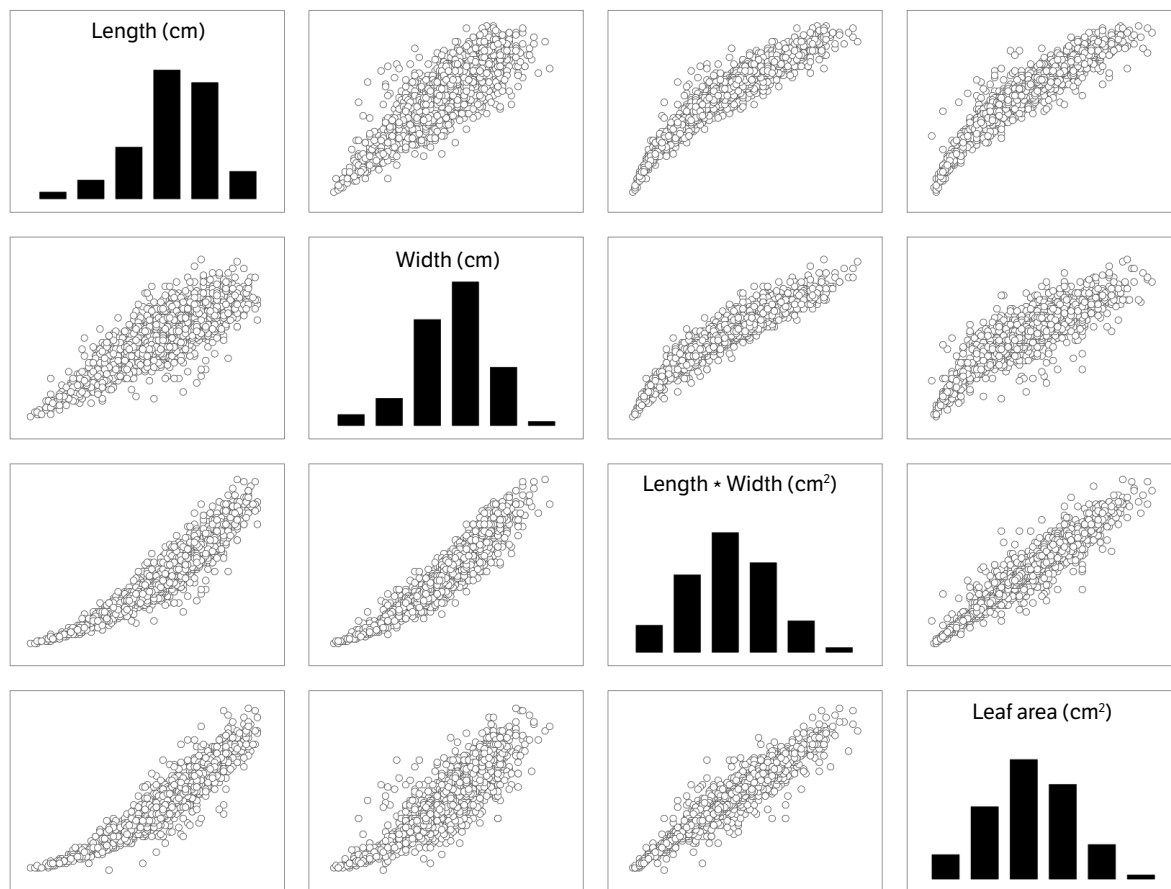
Figure 3. Matrix with histograms (diagonal) and scatter plots between length, width, product between length and width, and leaflet area of 2,605 leaflets (pooled data) used for the generation of models to estimate the peanut leaflet area.

Regarding the regression models obtained to estimate the peanut LA, it was observed that the determination coefficients (R2) and linear correlation coefficients (r) were in the ranges of 0.7397-0.9904 and 0.8601-0.9641 (IAC Caiapó), 0.8973-0.9984 and 0.9416-0.9925 (IAC 8112), 0.7630-0.9776 and 0.8735-0.9351 (Runner IAC 886), 0.8868-0.9975 and 0.9417-0.9890 (BRS Havana), 0.9286-0.9982 and 0.9636-0.9944 (BRS 151 L7), and 0.8968-0.9975 and 0.9470-0.9910 (IAC Tatuí) (Table 1). According to the criteria for choosing the best models, models #9 and #10 are not indicated to accurately estimate the LA of any of the cultivars, as they presented the lowest coefficients of determination, correlation coefficients linear, and Willmott's concordance index, and higher AIC and RMSE, confirming the ineffectiveness of these models (Table 1). For each cultivar individually, Eq. 7, obtained with the power model using the product between leaflet length and width (LW), was the most suitable for estimating the LA with greater precision (> 97%). However, other models can also be indicated to estimate the LA of each cultivar individually, such as models #3 and #4 (Table 1).

It was observed in the principal components analysis (PCA) that there was concentration of 98.01% data variability in the two main axes, corresponding to 69.94% in the first and 28.07% in the second (Figs. 4a and b). PCA showed clustering among cultivars, evidencing similar phylogenetic characteristics between them. This analysis indicated slight morphological variation between the peanut cultivars' leaflets, confirmed by the leaflet morphotypes of each cultivar (Fig. 4c). These results and the established criteria made it possible to create a generalized model covering all cultivars. The power model $\hat{y} = 0.875 * LW^{0.929}$ was the most suitable to estimate the peanut LA with an accuracy of 99%, regardless of cultivar.

Table 1. Statistical models, regression coefficients (β_0 and β_1), coefficient of determination (R^2), Pearson's linear correlation coefficient (r), Akaike information criterion (AIC), Willmott's agreement index (d), root mean square error (RMSE), and equations for estimating leaflet area of peanut cultivars as a function of linear leaflet dimensions (length and width).

| Equation code	Model	Coefficients		R^2	r	AIC	d	RMSE	Estimator of LA ($\hat{y}$)
		β_0	β_1						
				IAC Caiapó					
1	$\hat{y} = \beta_0 + \beta_1 * L + \varepsilon_i$	-5.528	3.127	0.8964	0.9469	1,148.85	0.9721	1.5018	$\hat{y} = -5.528 + 3.127 * L$
2	$\hat{y} = \beta_0 + \beta_1 * W + \varepsilon_i$	-6.473	6.545	0.7669	0.8761	1,402.66	0.9305	2.2527	$\hat{y} = -6.473 + 6.545 * W$
3	$\hat{y} = \beta_0 + \beta_1 * LW + \varepsilon_i$	0.735	0.677	0.9280	0.9634	1,034.75	0.9810	1.2516	$\hat{y} = 0.735 + 0.677 * LW$
4	$\hat{y} = \beta_1 * LW + \varepsilon_i$	---	0.715	0.9295	0.9634	1,046.82	0.9812	1.2800	$\hat{y} = 0.715 * LW$
5	$\hat{y} = \beta_0 * L^{\beta_1} + \varepsilon_i$	0.741	1.596	0.9073	0.9525	1,114.87	0.9750	1.4225	$\hat{y} = 0.741 * L^{1.596}$
6	$\hat{y} = \beta_0 * W^{\beta_1} + \varepsilon_i$	2.186	1.619	0.7683	0.8765	1,402.23	0.9293	2.2511	$\hat{y} = 2.186 * W^{1.619}$
7	$\hat{y} = \beta_0 * LW^{\beta_1} + \varepsilon_i$	0.914	0.918	0.9904	0.9641	1,029.15	0.9813	1.2404	$\hat{y} = 0.914 * LW^{0.918}$
8	$\hat{y} = \beta_0 * \beta_1^{L} + \varepsilon_i$	2.423	1.315	0.8884	0.9425	1,176.70	0.9682	1.5701	$\hat{y} = 2.423 * 1.315^{L}$
9	$\hat{y} = \beta_0 * \beta_1^{W} + \varepsilon_i$	2.470	1.721	0.7397	0.8601	1,440.54	0.9160	2.3932	$\hat{y} = 2.470 * 1.721^{W}$
10	$\hat{y} = \beta_0 * \beta_1^{LW} + \varepsilon_i$	4.943	1.052	0.7397	0.8601	1,233.43	0.9160	2.3932	$\hat{y} = 4.943 * 1.052^{LW}$
				IAC 8112					
1	$\hat{y} = \beta_0 + \beta_1 * L + \varepsilon_i$	-7.808	3.319	0.9489	0.9741	990.95	0.9867	0.9093	$\hat{y} = -7.808 + 3.319 * L$
2	$\hat{y} = \beta_0 + \beta_1 * W + \varepsilon_i$	-8.899	7.745	0.8863	0.9416	1,288.50	0.9691	1.3564	$\hat{y} = -8.899 + 7.745 * W$
3	$\hat{y} = \beta_0 + \beta_1 * LW + \varepsilon_i$	0.276	0.705	0.9850	0.9924	534.88	0.9962	0.4925	$\hat{y} = 0.276 + 0.705 * LW$
4	$\hat{y} = \beta_1 * LW + \varepsilon_i$	---	0.720	0.9851	0.9924	544.86	0.9961	0.5005	$\hat{y} = 0.720 * LW$
5	$\hat{y} = \beta_0 * L^{\beta_1} + \varepsilon_i$	0.562	1.701	0.9493	0.9743	989.92	0.9866	0.9080	$\hat{y} = 0.562 * L^{1.701}$
6	$\hat{y} = \beta_0 * W^{\beta_1} + \varepsilon_i$	1.696	1.941	0.9044	0.9510	1,224.97	0.9744	1.2454	$\hat{y} = 1.696 * W^{1.941}$
7	$\hat{y} = \beta_0 * LW^{\beta_1} + \varepsilon_i$	0.784	0.971	0.9984	0.9925	531.24	0.9962	0.4901	$\hat{y} = 0.784 * LW^{0.971}$
8	$\hat{y} = \beta_0 * \beta_1^{L} + \varepsilon_i$	2.285	1.310	0.9258	0.9622	1,136.19	0.9794	1.1053	$\hat{y} = 2.285 * 1.310^{L}$
9	$\hat{y} = \beta_0 * \beta_1^{W} + \varepsilon_i$	1.692	2.027	0.8973	0.9472	1,253.23	0.9715	1.2936	$\hat{y} = 1.692 * 2.027^{W}$
10	$\hat{y} = \beta_0 * \beta_1^{LW} + \varepsilon_i$	4.652	1.056	0.8973	0.9472	1,055.37	0.9715	1.2936	$\hat{y} = 4.652 * 1.056^{LW}$
				Runner IAC 886					
1	$\hat{y} = \beta_0 + \beta_1 * L + \varepsilon_i$	-4.988	3.038	0.8007	0.8951	1,630.58	0.9421	2.1232	$\hat{y} = -4.988 + 3.038 * L$
2	$\hat{y} = \beta_0 + \beta_1 * W + \varepsilon_i$	-5.189	5.800	0.7800	0.8835	1,667.54	0.9350	2.2308	$\hat{y} = -5.189 + 5.800 * W$
3	$\hat{y} = \beta_0 + \beta_1 * LW + \varepsilon_i$	1.080	0.640	0.8742	0.9351	1,458.52	0.9655	1.6869	$\hat{y} = 1.080 + 0.640 * LW$
4	$\hat{y} = \beta_1 * LW + \varepsilon_i$	---	0.700	0.8745	0.9351	1,482.79	0.9658	1.7472	$\hat{y} = 0.700 * LW$
5	$\hat{y} = \beta_0 * L^{\beta_1} + \varepsilon_i$	0.692	1.642	0.8207	0.9059	1,592.24	0.9491	2.0171	$\hat{y} = 0.692 * L^{1.642}$
6	$\hat{y} = \beta_0 * W^{\beta_1} + \varepsilon_i$	2.138	1.568	0.7885	0.8879	1,653.90	0.9374	2.1905	$\hat{y} = 2.138 * W^{1.568}$
7	$\hat{y} = \beta_0 * LW^{\beta_1} + \varepsilon_i$	0.974	0.888	0.9776	0.9351	1,458.47	0.9661	1.6868	$\hat{y} = 0.974 * LW^{0.888}$
8	$\hat{y} = \beta_0 * \beta_1^{L} + \varepsilon_i$	2.127	1.347	0.8121	0.9011	1,611.14	0.9441	2.0688	$\hat{y} = 2.127 * 1.347^{L}$
9	$\hat{y} = \beta_0 * \beta_1^{W} + \varepsilon_i$	2.432	1.679	0.7630	0.8735	1,699.11	0.9252	2.3270	$\hat{y} = 2.432 * 1.679^{W}$
10	$\hat{y} = \beta_0 * \beta_1^{LW} + \varepsilon_i$	4.704	1.052	0.7630	0.8735	1,598.30	0.9252	2.3270	$\hat{y} = 4.704 * 1.052^{LW}$
				BRS Havana					
1	$\hat{y} = \beta_0 + \beta_1 * L + \varepsilon_i$	-7.238	3.197	0.9170	0.9577	1,245.41	0.9779	1.1526	$\hat{y} = -7.238 + 3.197 * L$
2	$\hat{y} = \beta_0 + \beta_1 * W + \varepsilon_i$	-7.311	7.073	0.9084	0.9532	1,284.32	0.9755	1.2105	$\hat{y} = -7.311 + 7.073 * W$
3	$\hat{y} = \beta_0 + \beta_1 * LW + \varepsilon_i$	0.527	0.676	0.9774	0.9886	728.93	0.9942	0.6014	$\hat{y} = 0.527 + 0.676 * LW$
4	$\hat{y} = \beta_1 * LW + \varepsilon_i$	---	0.704	0.9782	0.9886	759.23	0.9940	0.6264	$\hat{y} = 0.704 * LW$
5	$\hat{y} = \beta_0 * L^{\beta_1} + \varepsilon_i$	0.456	1.806	0.9377	0.9683	1,131.88	0.9836	0.9990	$\hat{y} = 0.456 * L^{1.806}$
6	$\hat{y} = \beta_0 * W^{\beta_1} + \varepsilon_i$	2.114	1.707	0.9114	0.9546	1,273.23	0.9759	1.1937	$\hat{y} = 2.114 * W^{1.707}$
7	$\hat{y} = \beta_0 * LW^{\beta_1} + \varepsilon_i$	0.840	0.940	0.9975	0.9890	714.36	0.9944	0.5904	$\hat{y} = 0.840 * LW^{0.940}$
8	$\hat{y} = \beta_0 * \beta_1^{L} + \varepsilon_i$	1.952	1.341	0.9218	0.9601	1,227.91	0.9783	1.1274	$\hat{y} = 1.952 * 1.341^{L}$
9	$\hat{y} = \beta_0 * \beta_1^{W} + \varepsilon_i$	2.183	1.835	0.8868	0.9417	1,374.45	0.9676	1.3560	$\hat{y} = 2.183 * 1.835^{W}$
10	$\hat{y} = \beta_0 * \beta_1^{LW} + \varepsilon_i$	4.659	1.055	0.8868	0.9417	1,216.86	0.9676	1.3560	$\hat{y} = 4.659 * 1.055^{LW}$

Table 1. Continuation...

| Equation code	Model	Coefficients β_0	Coefficients β_1	R^2	r	AIC	d	RMSE	Estimator of LA ($\hat{y}$)
		BRS 151 L7							
1	$\hat{y} = \beta_0 + \beta_1 * L + \varepsilon_i$	-7.607	3.298	0.9305	0.9647	1,221.39	0.9817	1.3467	$\hat{y} = -7.607 + 3.298 * L$
2	$\hat{y} = \beta_0 + \beta_1 * W + \varepsilon_i$	-7.411	7.159	0.9366	0.9678	1,188.54	0.9834	1.2857	$\hat{y} = -7.411 + 7.159 * W$
3	$\hat{y} = \beta_0 + \beta_1 * LW + \varepsilon_i$	0.192	0.693	0.9880	0.9943	575.27	0.9971	0.5407	$\hat{y} = 0.192 + 0.693 * LW$
4	$\hat{y} = \beta_1 * LW + \varepsilon_i$	---	0.702	0.9888	0.9943	580.50	0.9971	0.5462	$\hat{y} = 0.702 * LW$
5	$\hat{y} = \beta_0 * L^{\beta_1} + \varepsilon_i$	0.366	1.928	0.9599	0.9797	1,027.09	0.9896	1.0235	$\hat{y} = 0.366 * L^{1.928}$
6	$\hat{y} = \beta_0 * W^{\beta_1} + \varepsilon_i$	1.884	1.812	0.9468	0.9730	1,127.54	0.9860	1.1795	$\hat{y} = 1.884 * W^{1.812}$
7	$\hat{y} = \beta_0 * LW^{\beta_1} + \varepsilon_i$	0.750	0.978	0.9982	0.9944	573.87	0.9971	0.5396	$\hat{y} = 0.750 * LW^{0.978}$
8	$\hat{y} = \beta_0 * \beta_1^{L} + \varepsilon_i$	1.743	1.363	0.9432	0.9711	1,160.53	0.9841	1.2358	$\hat{y} = 1.743 * 1.363^{L}$
9	$\hat{y} = \beta_0 * \beta_1^{W} + \varepsilon_i$	1.898	1.914	0.9286	0.9636	1,239.21	0.9800	1.3811	$\hat{y} = 1.898 * 1.914^{W}$
10	$\hat{y} = \beta_0 * \beta_1^{LW} + \varepsilon_i$	4.466	1.055	0.9286	0.9636	1,232.40	0.9800	1.3811	$\hat{y} = 4.466 * 1.055^{LW}$
		IAC Tatuí							
1	$\hat{y} = \beta_0 + \beta_1 * L + \varepsilon_i$	-6.688	3.147	0.9260	0.9623	1,295.45	0.9804	1.2127	$\hat{y} = -6.688 + 3.147 * L$
2	$\hat{y} = \beta_0 + \beta_1 * W + \varepsilon_i$	-8.659	7.411	0.9086	0.9533	1,379.91	0.9755	1.3477	$\hat{y} = -8.659 + 7.411 * W$
3	$\hat{y} = \beta_0 + \beta_1 * LW + \varepsilon_i$	0.164	0.692	0.9821	0.9910	727.08	0.9954	0.5959	$\hat{y} = 0.164 + 0.692 * LW$
4	$\hat{y} = \beta_1 * LW + \varepsilon_i$	---	0.701	0.9822	0.9910	729.29	0.9954	0.5991	$\hat{y} = 0.701 * LW$
5	$\hat{y} = \beta_0 * L^{\beta_1} + \varepsilon_i$	0.493	1.774	0.9447	0.9719	1,179.73	0.9856	1.0494	$\hat{y} = 0.493 * L^{1.774}$
6	$\hat{y} = \beta_0 * W^{\beta_1} + \varepsilon_i$	1.635	1.914	0.9213	0.9598	1,321.40	0.9788	1.2527	$\hat{y} = 1.635 * W^{1.914}$
7	$\hat{y} = \beta_0 * LW^{\beta_1} + \varepsilon_i$	0.739	0.982	0.9975	0.9910	726.21	0.9954	0.5953	$\hat{y} = 0.739 * LW^{0.982}$
8	$\hat{y} = \beta_0 * \beta_1^{L} + \varepsilon_i$	1.942	1.344	0.9306	0.9646	1,276.71	0.9808	1.1846	$\hat{y} = 1.942 * 1.344^{L}$
9	$\hat{y} = \beta_0 * \beta_1^{W} + \varepsilon_i$	1.790	1.937	0.8968	0.9470	1,436.96	0.9702	1.4474	$\hat{y} = 1.790 * 1.937^{W}$
10	$\hat{y} = \beta_0 * \beta_1^{LW} + \varepsilon_i$	4.380	1.057	0.8968	0.9470	1,284.08	0.9702	1.4474	$\hat{y} = 4.380 * 1.057^{LW}$
		Pooled data							
1	$\hat{y} = \beta_0 + \beta_1 * L + \varepsilon_i$	-6.200	3.114	0.8709	0.9332	9,980.91	0.9644	1.6427	$\hat{y} = -6.200 + 3.114 * L$
2	$\hat{y} = \beta_0 + \beta_1 * W + \varepsilon_i$	-5.859	6.411	0.7771	0.8815	11,403.4	0.9341	2.1587	$\hat{y} = -5.859 + 6.411 * W$
3	$\hat{y} = \beta_0 + \beta_1 * LW + \varepsilon_i$	0.660	0.674	0.9183	0.9583	8,789.13	0.9783	1.3067	$\hat{y} = 0.660 + 0.674 * LW$
4	$\hat{y} = \beta_1 * LW + \varepsilon_i$	---	0.709	0.9189	0.9583	8,880.40	0.9786	1.3303	$\hat{y} = 0.709 * LW$
5	$\hat{y} = \beta_0 * L^{\beta_1} + \varepsilon_i$	0.564	1.714	0.8882	0.9424	9,606.55	0.9696	1.5288	$\hat{y} = 0.564 * L^{1.714}$
6	$\hat{y} = \beta_0 * W^{\beta_1} + \varepsilon_i$	2.297	1.593	0.7806	0.8835	11,363.9	0.9345	2.1423	$\hat{y} = 2.297 * W^{1.593}$
7	$\hat{y} = \beta_0 * LW^{\beta_1} + \varepsilon_i$	0.875	0.929	0.9883	0.9586	8,770.83	0.9784	1.3021	$\hat{y} = 0.875 * LW^{0.929}$
8	$\hat{y} = \beta_0 * \beta_1^{L} + \varepsilon_i$	2.107	1.332	0.8723	0.9340	9,976.25	0.9633	1.6412	$\hat{y} = 2.107 * 1.332^{L}$
9	$\hat{y} = \beta_0 * \beta_1^{W} + \varepsilon_i$	2.456	1.735	0.7544	0.8685	11,675.9	0.9219	2.2746	$\hat{y} = 2.456 * 1.735^{W}$
10	$\hat{y} = \beta_0 * \beta_1^{LW} + \varepsilon_i$	4.700	1.054	0.7544	0.8685	10,184.6	0.9219	2.2746	$\hat{y} = 4.700 * 1.054^{LW}$

The equation proposed to estimate the area of peanut leaflets presented high adjustment (R^2 = 0.9883) of the data about the adjustment of the line, in which the residual variance was homogeneous with low dispersion of the data (normal distribution) (Fig. 5). The area of leaflets estimated from the equation recommended in the present study showed positive correlation with the observed area of leaflets (determined by digital images), with the coefficient of determination (R^2) of 0.9791, confirming a significant relationship between these parameters (Figs. 6a and 6b). Thus, regardless of the cultivar, the equation $\hat{y} = 0.875 * LW^{0.929}$ can efficiently estimate the area of peanut leaflets by speed using product dimensions between LW.
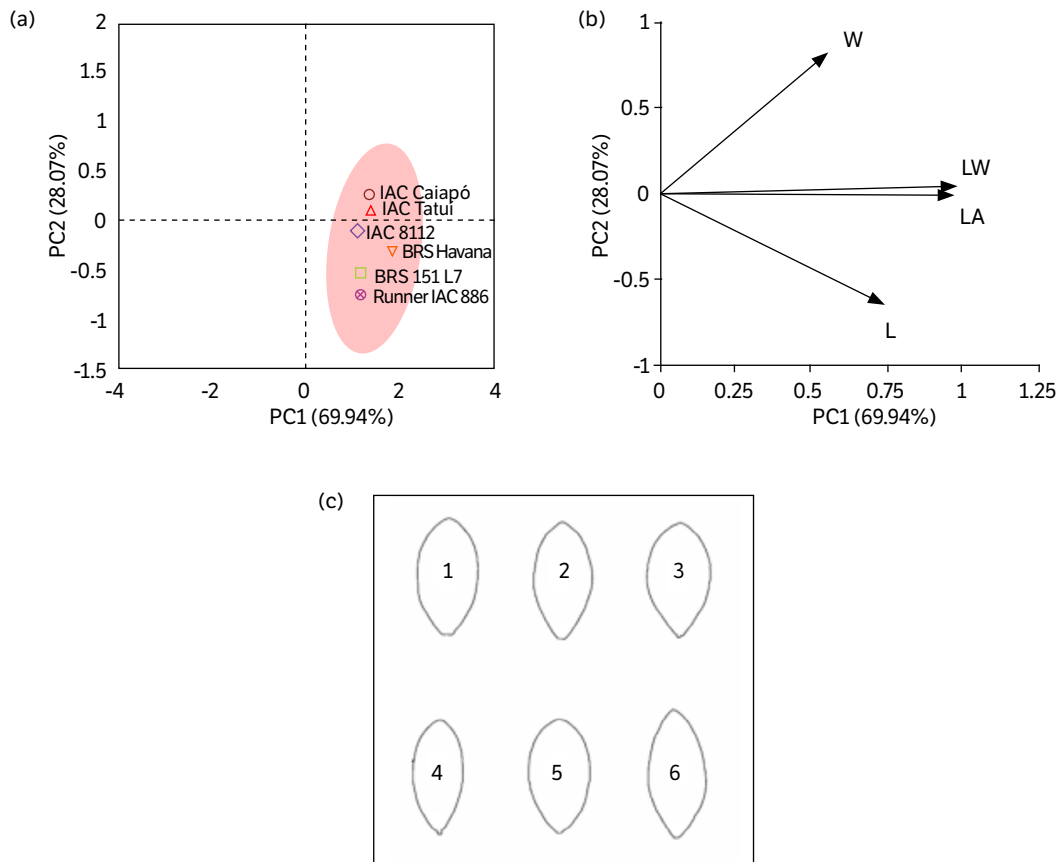
PC: principal component; W: width; L: length; LW: product between length and width; LA: leaflet area.

Figure 4. (a) Principal component analysis of six peanut cultivars leaf parameters. The large circle represents the cluster formed by the Euclidean similarity distance. (b) Loading plot between length, width, product between length and width, and leaflet area of peanut cultivars. (c) Design of peanut cultivars: IAC Caiapó (1), IAC 8112 (2), Runner IAC 886 (3), BRS Havana (4), BRS 151 L7 (5), and IAC Tatuí (6).
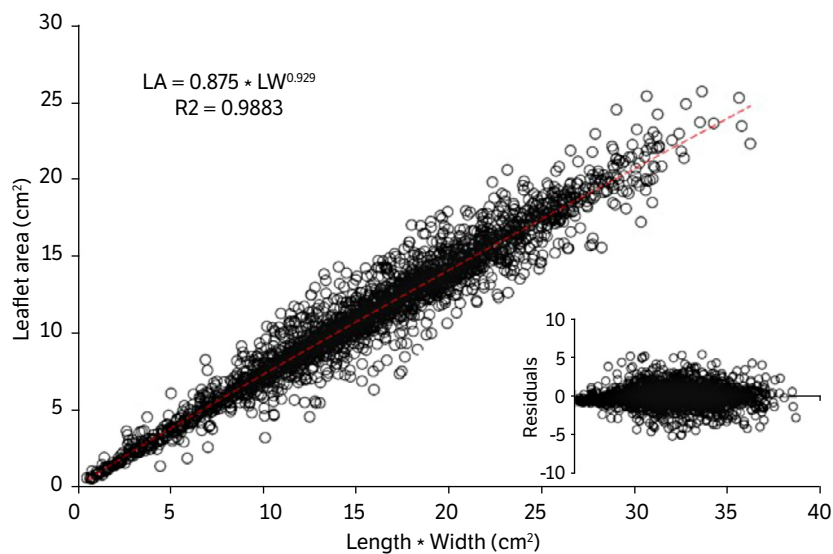
LA: leaflet area; L: length; LW: product between length and width.

Figure 5. Relation between the observed leaflet area and the product between length and width of leaflets of peanut cultivars (pooled data), from the model LA = 0.875*LW0.929. The analysis of the dispersion pattern of the residues is presented in the insert.

(a)

$\hat{y} = 0.6521 + 0.9511x$

$R^2 = 0.9791$

(b)

$t = 0.5033; p = 0.6148^{ns}$
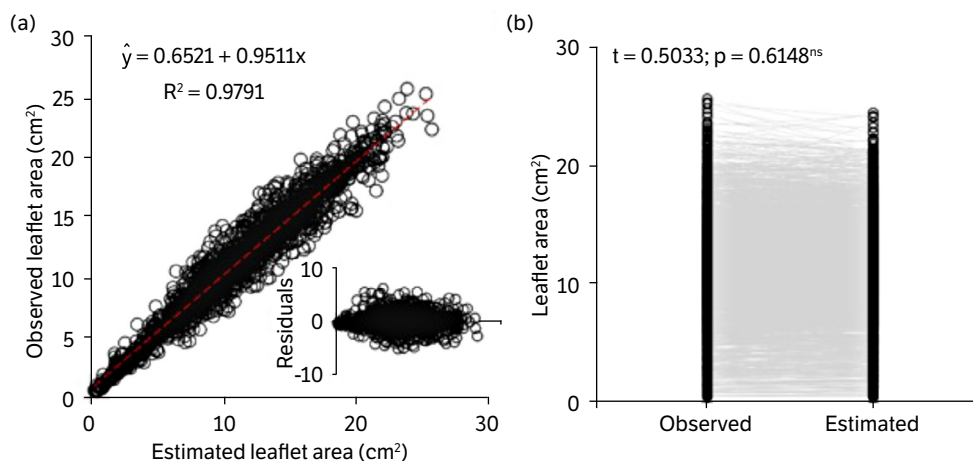
Figure 6. (a) Relationship between the observed leaflet area and the leaflet area estimated by the model as a function of the product of length and width (LA = 0.875*LW0.929). The analysis of the dispersion pattern of the residues is presented in the insert. (b) Observed and estimated leaflet areas compared with each other using Student's t-test (p < 0.01).

DISCUSSION

This study describes a non-destructive method to estimate the LA of peanut cultivars, seeking to build a single model that allows estimating the LA of six cultivars. Therefore, we observed that the cultivars showed little intraspecific difference in leaflet morphology, confirming the possibility of building a single model to estimate the peanut LA.

The wide variability observed in the data of LW and LA (CV > 33.9%) is fundamental for studies related to allometric modeling to estimate the leaf area of plant species (Dalmago et al. 2019, Toebe et al. 2019). This high variability allows for more representative models and precise equations for leaves of different shapes and sizes, which can measure at different phenological stages during the plant cycle (Cargnelutti Filho et al. 2021). Although the implemented experiment was in only one area, it is considered that the total number of leaflets sampled (2,605) in all parts of the plants (middle, lower and upper thirds) was adequate for the construction of models that determine the area of peanut leaflets as a function of linear dimensions (L, W and or LW). Some studies have reported that a small number of samples (leaves or leaflets) to build allometric models can generate biased and unreliable equations for estimating the leaf area of plants (Pompelli et al. 2012).

The present study showed that the best equations to estimate the LA of peanut cultivars were those that used LW, compared to equations that used only one leaflet dimension (L or W), with the best criteria and adjustments of the models used (Guimarães et al. 2019, Macário et al. 2020, Goergen et al. 2021), except for the exponential model, in which the best equations were those that used L of the leaflets (Ribeiro et al. 2020).

Leaf area can also be estimated using only a single dimension of the leaf surface (L or W), simplifying the analyses. However, using equations with only one size can lead to a loss of model precision for the leaf area estimation (Pompelli et al. 2012), which was observed in the present study with the minor adjustments of the proposed models, except for the exponential model, in which the equation using leaflet L was the most adequate to estimate LA.

Generally, linear models are the most used to estimate the leaf area of agricultural and forest species (Gomes et al. 2020, Goergen et al. 2021, Hernandéz-Fernandez et al. 2021, Mela et al. 2022). However, it is assumed that these models are used with losses in precision, which often occurs due to the high values of the intercepts (β_0) of the regression line (Santos et al. 2021). Therefore, in this study, we observed that we disregarded the use of linear models #1 and #2 for all peanut cultivars. According to Zuur et al. (2010), linear models show high residual dispersion, indicating that these models are not recommended to estimate leaf area during leaf development throughout the crop cycle, whereas nonlinear models are the most suitable for determining leaf area. The nonlinear Eq. 8, obtained through the power model, was the most appropriate

to estimate the LA of the cultivars individually and in general. This model was also the most suitable for estimating the leaf area of other plant species (Trachta et al. 2020, Toebe et al. 2021).

The proposed equations can estimate the LA of peanut cultivars at different phenological stages and crop planting conditions. Thus, the data obtained are fundamental for researchers who seek to evaluate this species vegetative growth and development, physiology, and production, in different environments and planting conditions.

CONCLUSION

We can apply the non-destructive method using allometric models accurately to estimate the peanut LA from the linear dimensions of the leaflet. Peanut LA can be more accurately estimated using the equation $\hat{y} = 0.875 * LW^{0.929}$ using LW. This equation will allow researchers to make non-destructive measurements or repeat measurements on the same leaves or plants during the crop cycle.

AUTHORS' CONTRIBUTION

Conceptualization: Ribeiro J. E. S. and Coêlho E. S.; **Formal Analysis:** Oliveira P. H. A., Lopes W. A. R., Silva E. F. and Oliveira A. K. S.; **Funding Acquisition:** Ribeiro J. E. S., Silveira L. M., Silva D. V. and Barros Júnior A. P.; **Investigation:** Ribeiro J. E. S.; **Methodology:** Ribeiro J. E. S., Coêlho E. S., Dias T. J. and Lopes W. A. R; **Project Administration:** Ribeiro J. E. S.; **Resources:** Ribeiro J. E. S.; **Supervision:** Ribeiro J. E. S.; **Validation:** Ribeiro J. E. S.; **Visualization:** Ribeiro J. E. S., Silveira L. M., Silva D. V., Barros Júnior A. P. and Dias T. J.; **Writing – Original Draft:** Ribeiro J. E. S.; **Writing – Review and Editing:** Ribeiro J. E. S., Silveira L. M., Silva D. V., Barros Júnior A. P. and Dias T. J..

DATA AVAILABILITY STATEMENT

All dataset were generated and analyzed in the current study.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Akram, N. A., Shafiq, F. and Ashraf, M. (2018). Peanut (*Arachis hypogaea* L.): A prospective legume crop to offer multiple health benefits under changing climate. Comprehensive Reviews in Food Science and Food Safety, 17, 1325-1338. https://doi.org/10.1111/1541-4337.12383

Alvares, C. A., Stape, J. L., Sentelhas, P. C., Gonçalves, J. L. M. and Sparovek, G. (2013). Köppen's climate classification map for Brazil. Meteorologische Zeitschrift, 22, 711-728. https://doi.org/10.1127/0941-2948/2013/0507

Cargnelutti Filho, A., Pezzini, R. V., Neu, I. M. M. and Dumke, G. E. (2021). Estimation of buckwheat leaf area by leaf dimensions. Semina, 42, 1529-1548. https://doi.org/10.5433/1679-0359.2021v42n3Supl1p1529

Córcoles, A. D., Magesan, E., Srinivasan, S. J., Cross, A. W., Steffen, M., Gambetta, J. M. and Chow, J. M. (2015). Demonstration of a quantum error detection code using a square lattice of four superconducting qubits. Nature Communications, 6, 6979. https://doi.org/10.1038/ncomms7979

Dalmago, G. A., Bianchi, C. A. M., Kovaleski, S. and Fochesatto, E. (2019). Evaluation of mathematical equations for estimating leaf area in rapeseed. Ciência Agronômica, 50, 420-430. https://doi.org/10.5935/1806-6690.20190050

[Embrapa] Empresa Brasileira de Pesquisa Agropecuária. (2018). Sistema Brasileiro de classificação de solos. Brasília: Embrapa Solos.

Ghoreishi, M., Hossini, Y. and Maftoon, M. (2012). Simple models for predicting leaf area of mango (*Mangifera indica* L.). Journal of Biology and Earth Sciences, 2, B45-B53.

Gill, J. (1986). Outliers, residuals, and influence in multiple regression. Journal of Animal Breeding and Genetics, 103, 161-175. https://doi.org/10.1111/j.1439-0388.1986.tb00079.x

Goergen, P. C. H., Lago, I., Schwab, N. T., Alves, A. F., Freitas, C. P. O., Verlaine, S. and Selli, V. S. (2021). Allometric relationship and leaf area modeling estimation on chia by nondestructive method. Brazilian Journal of Agricultural and Environmental Engineering, 25, 305-311. https://doi.org/10.1590/1807-1929/agriambi.v25n5p305-311

Gomes, F. R., Silva, D. F. P., Ragagnin, A. L. S. L., Souza, P. H. M. and Cruz, S. C. S. (2020). Leaf area estimation of *Anacardium humile*. Revista Brasileira de Fruticultura, 42, 1-8. https://doi.org/10.1590/0100-29452020628

Guimarães, M. J. M., Coelho Filho, M. A., Gomes Junior, F. A., Silva, M. A. M., Alves, C. V. O. and Lopes, I. (2019). Modelos matemáticos para a estimativa da área foliar de mandioca. Revista de Ciências Agrárias, 62, 1-5. https://doi.org/10.22491/rca.2019.3015

Hernandéz-Fernandez, I. A., Jarma-Orozco, A. and Pompelli, M. F. (2021). Allometric models for non-destructive leaf area measurement of stevia: An in depth and complete analysis. Horticultura Brasileira, 39, 205-215. https://doi.org/10.1590/s0102-0536-20210212

Keramatlou, I., Sharifani, M., Sabouri, H., Alizadeh, M, and Kamkar, B. (2015). A simple linear model for leaf area estimation in Persian walnut (*Juglans regia* L.). Scientia Horticulturae, 184, 36-39. https://doi.org/10.1016/j.scienta.2014.12.017

Krishna, G., Singh, B. K., Kim, E. K., Morya, V. K. and Ramteke, P. W. (2015). Progress in genetic engineering of peanut (*Arachis hypogaea* L.): a review. Plant Biotechnology Journal, 13, 147-162. https://doi.org/10.1111/pbi.12339

Liu, S., Baret, F., Abichou, M., Manceau, L., Andrieu, B., Weiss, M. and Martre, P. (2021). Importance of the description of light interception in crop growth models. Plant Physiology, 186, 977-997. https://doi.org/10.1093/plphys/kiab113

Lucena, L. R. R., Leite, M. L. M. V., Simões, V. J. L. P., Nóbrega, C., Almeida, M. C. R. A. and Simplicio, J. B. (2021). Estimating the área and weight of cactos forage cladodes using linear dimensions. Acta Scientiarum. Agronomy, 43, e45460. https://doi.org/10.4025/actasciagron.v43i1.45460

Macário, A. P. S., Ferraz, R. L. S., Costa, P. S., Brito Neto, J. F., Melo, A. S. and Dantas Neto, J. (2020). Allometric models of estimating *Moringa oleífera* leaflets area. Ciência e Agrotecnologia, 44, 1-10. https://doi.org/10.1590/1413-7054202044005220

Marquardt, D. W. (1970). Generalized inverse, ridge regression, biased linear estimation and nonlinear estimation. Technometrics, 12, 591-612. https://doi.org/10.2307/1267205

Mattos, E. M., Binkley, D., Campoe O. C., Alvares, C. A. and Stape, J. L. (2020). Variation in canopy structure, leaf area, light interception and light use efficiency among Eucalyptus clones. Forest Ecology and Management, 463, 118038. https://doi.org/10.1016/j.foreco.2020.118038

Mela, D., Dias, M. G., Silva, T. I., Ribeiro, J. E. S., Martinez, A. C. P. and Zuin, A. H. L. (2022). Estimation of *Thunbergia grandiflora* leaf area from allometric models. Comunicata Scientiae, 13, e3722. https://doi.org/10.14295/cs.v13.3722

Peksen, E. (2007). Non-destructive leaf area estimation model for faba bean (*Vicia faba* L.). Scientia Horticulturae, 113, 322-328. https://doi.org/10.1016/j.scienta.2007.04.003

Pompelli, M. F., Antunes, W. C., Ferreira, D. T. R. G., Cavalcante, P. P. G. S., Wanderley-Filho, H. C. L. and Endres, L. (2012). Allometric models for non-destructive leaf area estimation of the *Jatropha curcas*. Biomass & Bioenergy, 36, 77-85. https://doi.org/10.1016/j.biombioe.2011.10.010

R Core Team. (2022). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Ribeiro, J. E. S., Coêlho, E. S., Figueiredo, F. R. A. and Melo, M. F. (2020). Non-destructive method for estimating leaf area of *Erythroxylum pauferrense* (Erythroxylaceae) from linear dimensions of leaf blades. Acta Botanica Mexicana, 127, e1717. https://doi.org/10.21829/abm127.2020.1717

Sabouri, H. and Sajadi, S. (2022). Image processing and area estimation of chia (*Salvia hispanica* L.), quinoa (*Chenopodium quinoa* Willd.), and bitter melon (*Momordica charantia* L.) leaves based on statistical and intelligent methods. Journal of Applied Research on Medicinal and Aromatic Plants, 30, 100382. https://doi.org/10.1016/j.jarmap.2022.100382

Salazar, J. C. S., Muñoz, L. M. M., Bautista, E. H. D., Rienzo, J. A. D. and Casanoves, F. (2018). Non-destructive estimation of the leaf weight and leaf area in cacao (*Theobroma cacao* L.). Scientia Horticulturae, 229, 19-24. https://doi.org/10.1016/j.scienta.2017.10.034

Santos, J. N. B., Jarma-Orozco, A., Antunes, W. C., Mendes, K. R., Figueiroa, J. M., Pessoa, J. M. and Pompelli, M. F. (2021). New approaches to predict leaf area in woody tree species from the Atlantic Rainforest, Brazil. Austral Ecology, 46, 613-626. https://doi.org/10.1111/aec.13017

Sauceda-Acosta, C. P., Gonzalez-Hernandez, V. A., Sanchez-Soto, B. H., Sauceda-Acosta, R. H., Ramirez-Tobias, H. M. and Quintana-Quiroz, J. G. (2017). MACF-IJ, automated method for measuring color and leaf area through digital images. Agrociencia, 51, 409-423.

Toebe, M., Melo, P. J., Souza, R. R., Mello, A. C., and Tartaglia F. L. (2019). Leaf area estimation in triticale by leaf dimensions. Revista Brasileira de Ciências Agrárias, 14, e5656. https://doi.org/10.5039/agraria.v14i2a5656

Toebe, M., Soldateli, F. J., Souza, R. R., Mello, A. C. and Segatto, A. (2021). Leaf area estimation of *Burley tobacco*. Ciência Rural, 51, e20200071. https://doi.org/10.1590/0103-8478cr20200071

Toomer, O. T. (2017). Nutritional chemistry of the peanut (*Arachis hypogaea*). Critical Reviews in Food Science and Nutrition, 58, 3042-3053. https://doi.org/10.1080/10408398.2017.1339015

Trachta, M. A., Zanon, A. J., Alves, A. F., Freitas, C. P. O., Streck, N. A., Cardoso, P. S. and Rodrigues, L. B. (2020). Leaf area estimation with nondestructive method in cassava. Bragantia, 79, 472-484. https://doi.org/10.1590/1678-4499.20200018

Yeshitila, M. and Taye, M. (2016). Non-destructive prediction models for estimation of leaf area for most commonly grown vegetable crops in ethiopia. Science Journal of Applied Mathematics and Statistics, 4, 202-216. https://doi.org/10.11648/j.sjams.20160405.13

Zar, J. H. (1996). Biostatistical Analysis. Engewood Cliff: Prentice Hall.

Zhang, W. (2020). Digital image processing method for estimating leaf length and width tested using kiwifruit leaves (*Actinidia chinensis* Planch). PloS One, 15, e0235499. https://doi.org/10.1371/journal.pone.0235499

Zuur, A. F., Ieno, E. N. and Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. Methods in Ecology & Evolution, 1, 3-14. https://doi.org/10.1111/j.2041-210X.2009.00001.x