



AJUSTE DE MODELOS DE REGRESSÃO PARA ESTIMAR A ÁREA DE FOLHAS DE CARAMBOLA

Vinicius de Souza Oliveira¹, Gleyce Pereira Santos¹, Karina Tiemi Hassuda dos Santos¹, Hérica Chisté¹, Adriel Lima Nascimento², Basílio Cerri Neto³, Omar Schmildt¹, Sara Dousseau Arantes⁴, Robson Prucoli Posse⁵ and Edilson Romais Schmildt¹

¹Universidade Federal do Espírito Santo, Centro Universitário Norte do Espírito Santo, São Mateus, Espírito Santo, Brasil

²Universidade Federal do Espírito Santo, Alegre, Espírito Santo, Brasil

³Universidade Federal do Espírito Santo, Vitória, Espírito Santo, Brasil

⁴Instituto Capixaba de Pesquisa, Assistência Técnica e Extensão Rural (INCAPER), Linhares, Espírito Santo, Brasil

⁵Instituto Federal do Espírito Santo, Campus Itapina, Colatina, Espírito Santo, Brasil

ARTICLE INFO

Article History:

Received 29th June, 2019

Received in revised form

17th July, 2019

Accepted 20th August, 2019

Published online 28th September, 2019

Key Words:

Averrhoa carambola L.,
Método não Destrutivo,
Equações Matemáticas.

*Corresponding author:

ABSTRACT

O objetivo deste trabalho foi ajustar modelos de regressão matemática para estimar a área foliar de carambola (*Averrhoa carambola* L.) a partir das dimensões lineares da superfície foliar. Ao todo, 400 folhas foram digitalizadas e processadas através do software ImageJ[®], sendo obtidas as medidas do comprimento (C) ao longo da nervura central, maior largura (L), produto da multiplicação do comprimento com a largura (CL) e a área foliar observada (AFO). Para o ajuste da modelagem, foram utilizadas 350 folhas, onde a AFO utilizada como variável dependente em função de C, L ou CL, obtendo os modelos lineares de primeiro e segundo grau e modelo não linear do tipo potência. Com amostra de 50 folha, os modelos foram validados. Para isso, os valores de C, L e CL foram substituídos nas equações ajustadas na modelagem obtendo assim a área foliar estima (AFE). A AFE foi relacionada com a AFO através de um modelo de equação linear simples. A partir do teste t de student a 5% de probabilidade as hipóteses $H_0: \hat{\beta}_0 = 0$ versus $H_a: \hat{\beta}_0 \neq 0$ e $H_0: \hat{\beta}_1 = 1$ versus $H_a: \hat{\beta}_1 \neq 1$, foram testadas. Calculou-se ainda o erro absoluto médio (EAM), raiz do quadrado médio do erro (RQME) e índice de Willmott (d) para todas as equações. Assim, o modelo linear de segundo grau representado pela equação $AFE = -0,2712602 + 0,8324138(CL) - 0,0019955(CL)^2$, obtida a partir do produto da multiplicação do comprimento com a largura, foi o mais adequado para estimar a área foliar de carambola de maneira não destrutiva.

Copyright © 2019, Vanitha and Kalpanapriya. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Citation: Vanitha, R. and Kalpanapriya, D. 2019. "Ajuste de modelos de regressão para estimar a área de folhas de carambola", *International Journal of Development Research*, 09, (09), 29649-29652.

INTRODUCTION

Originária da Ásia, a caramboleira (*Averrhoa carambola* L.) pertence a família das Oxalidaceae, sendo uma espécie típica das regiões tropicais, no Brasil, é cultivada em todas as regiões, especialmente nas localidades com clima quente, sem a ocorrência de geadas (Bastos, 2004). O seu consumo varia nas mais diversas formas, sendo o fruto consumido *in natura* ou processado na forma de suco, geleia, doces e saladas, o extrato das sementes é utilizado para a remoção de manchas e as folhas fazem parte medicina popular (Andrade, Martins, 2007).

O estudo da área foliar é de extrema importância por esta relacionada ao aproveitamento da energia luminosa e consequentemente a capacidade fotossintética da planta (Gonzalez-Sanpedro *et al.*, 2008), permitindo estimar os padrões de crescimento e desenvolvimento da planta e consequentemente a sua produtividade total (Reis *et al.*, 2013). Modelos matemáticos para a estimativa da área foliar são obtidos através da análise de regressão, onde, a área das folhas é correlacionada com as medidas de seu comprimento e a largura (Astegiano *et al.*, 2001). Principalmente devido a seu

caráter não destrutivo, que permite a estimativa da área foliar em vários estádios de uma mesma planta durante todo seu ciclo de desenvolvimento, este método, tem sido amplamente utilizado por diversos autores para espécies frutíferas como *Vitis vinifera* L. (Buttaro et al., 2015), *Litchi chinensis* Sonn. (Oliveira et al., 2017), *Annona cherimila* Mill. (Schmidt et al., 2017), Guava (Vitória et al., 2018), *Artocarpus heterophyllus* (Oliveira et al., 2019). Entretanto, não se tem conhecimento de estudos que envolvam *Averrhoa carambola* L. Assim, devido a relevante importância da determinação da área foliar para estudos como parâmetros de crescimento vegetal, objetivou-se por este trabalho ajustar modelos de regressão para estimar com precisão a área das folhas de *Averrhoa carambola* L. a partir das dimensões lineares da superfície foliar.

METODOLOGIA

No presente estudo, foram utilizadas 400 folhas de carambola (*Averrhoa carambola* L.) coletadas no município de Linhares, localizado ao Norte do Estado do Espírito Santo, Brasil e situado a 19° 23' 28" de latitude sul, 40° 04' 20" de longitude e altitude 33 de metros. A precipitação média anual é de 1193 mm e temperatura média anual de 23,4 °C (Silva et al., 2016). O clima da região segundo a classificação de Köppen é do tipo *Aw* (tropical úmido) com chuvas durante o verão e inverno seco (Alvares et al., 2014). Após remoção do pecíolo, as folhas foram digitalizadas com scanner de mesa modelo Deskjet F2540® salvas com resolução de 75 dpi no formato Tag Image File Format (TIFF). As imagens foram processadas através do software ImageJ® (Schindelin et al., 2015), sendo feitas as medições do comprimento (C, em cm) ao longo da nervura central, maior largura (L, em cm) (Fig. 1) e a área foliar observada (AFO em cm²) de todas as folhas. Obteve-se também o produto da multiplicação do comprimento com a largura (CL, em cm²). Para todas estas variáveis, foi feito o estudo da estatística descritiva determinado os valores mínimos, máximos, médios, amplitude e coeficiente de variação (CV).

Para o ajuste dos modelos, com base na amostragem de 350 folhas, foram realizadas análises de regressão dos modelos linear de primeiro grau representada por $AFE = \hat{\beta}_0 + \hat{\beta}_1 x$, linear de segundo grau representada por $AFE = \hat{\beta}_0 + \hat{\beta}_1 x + \hat{\beta}_2 x^2$ e potencia representa por $AFE = \hat{\beta}_0 x^{\hat{\beta}_1}$. Onde a variável dependente foi a AFO sendo relacionada com as variáveis independentes C, L ou CL. Foi determinado também o coeficiente de determinação (R^2) para todos os modelos ajustados. Os valores dos parâmetros $\hat{\beta}_0$ (coeficiente linear), $\hat{\beta}_1$ (coeficiente angular) e $\hat{\beta}_2$ (coeficiente quadrático) foram estimados através do método de mínimos quadrados. Os modelos foram validados a partir de uma amostra com 50 folhas, em que os valores de C, L e CL foram substituídos nas equações ajustadas na modelagem obtendo assim a área foliar estimada (AFE) em cm² por cada modelo. Foi ajustado um modelo de equação linear simples ($AFE = \hat{\beta}_0 + \hat{\beta}_1 x$) com a AFE em função da AFO. As hipóteses $H_0: \hat{\beta}_0 = 0$ versus $H_a: \hat{\beta}_0 \neq 0$ e $H_0: \hat{\beta}_1 = 1$ versus $H_a: \hat{\beta}_1 \neq 1$, foram testadas por meio do teste t de Student a 5% de probabilidade. Para todas as equações efetuou-se cálculos para determinar o erro absoluto médio (EAM), raiz do quadrado médio do erro (RQME) e índice de Willmott (d) (Willmott, 1981), através das expressões 1, 2, 3.

$$EAM = \frac{\sum_{i=1}^n |AFE - AFO|}{n} \quad (1)$$

$$RQME = \sqrt{\frac{\sum_{i=1}^n (AFE - AFO)^2}{n}} \quad (2)$$

$$d = 1 - \left[\frac{\sum_{i=1}^n (AFE - AFO)^2}{\sum_{i=1}^n (|AFE - AFO| + |AFO - AFO|)^2} \right] \quad (3)$$

Onde: AFE são os valores estimado de área foliar; AFO são os valores observado de área foliar; \overline{AFO} é a média dos valores da área foliar observada; e n é o número de folhas amostradas usadas para a validação, sendo n=50 no presente estudo. A decisão para a seleção do melhor modelo que estime a área foliar de carambola foi baseada no coeficiente linear ($\hat{\beta}_0$) não diferente de zero, coeficiente angular ($\hat{\beta}_1$) não diferente de um, E, EAM e RQME mais próximos de zero e índice d de Willmott (Willmott, 1981) mais próximo de um. As análises estatísticas foram realizadas com o auxílio do software R (R Core Team, 2019), através de comando desenvolvidos para o pacote ExpDes.pt versão 1.2 (Ferreira et al., 2018).

RESULTADOS E DISCUSSÃO

Pela Tabela 1, observa-se que há elevada variação dos valores do comprimento (C), largura (L), produto do comprimento com a largura (CL) e área foliar observada (AFO). Essa elevada variabilidade dos dados amostras também foi relatada por outros autores (Oliveira et al., 2017; Oliveira et al., 2019) constatando que estes valores são desejáveis, uma vez que indicam folhas de tamanhos diferentes, mostrando representatividade dos dados amostrais em relação a população.

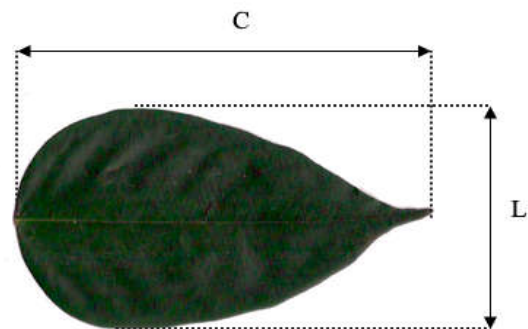


Fig. 1. Representação da medida do comprimento (C) ao longo da nervura principal e largura máxima (L) de folhas de *Averrhoa carambola* L.

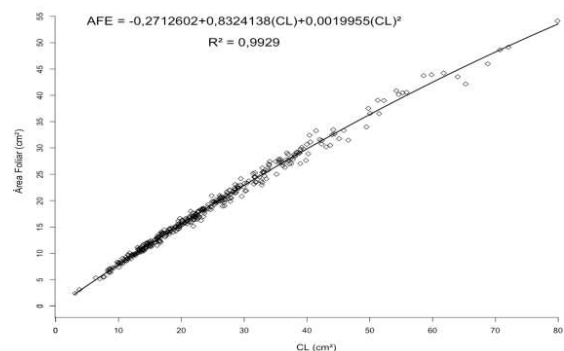


Fig. 3. Equação de modelo linear de segundo grau e coeficiente de determinação (R^2), utilizando a área foliar observada (AFO) como variável dependente, em função do produto do comprimento com a largura (CL) ajustada para estimar a área de folhas de *Averrhoa carambola* L., além do diagrama de dispersão do resíduo

Tabela 1. Análise descritiva com valor mínimo, máximo, médio, amplitude, e coeficiente de variação (CV) das variáveis: comprimento (C); largura (L); produto do comprimento com a largura (CL) e área foliar observada (AFO) de folhas *Averrhoa carambola* L. usadas para a modelagem e para a validação

Variável	Unidade	Mínimo	Máximo	Média	Amplitude	CV
350 folhas usadas para a modelagem						
L	cm	2,08	12,61	6,75	10,53	29,41
W	cm	1,49	6,65	3,47	5,16	21,10
LW	cm ²	3,11	79,89	24,75	76,78	50,48
OLA	cm ²	2,39	54,13	18,80	51,74	46,87
50 folhas usadas para a validação						
L	cm	3,03	10,68	6,22	7,65	30,86
W	cm	1,85	4,84	3,14	2,99	24,06
LW	cm ²	6,14	51,63	20,86	45,49	54,21
OLA	cm ²	4,88	37,10	16,05	32,22	51,52

Tabela 2. Equações de modelo linear de primeiro e segundo grau e potencia utilizando a área foliar observada (AFO) em função do comprimento (C), largura (L) e produto do comprimento com a largura (CL) e seus respectivos coeficientes de determinação (R²) para estimar a área de folhas de *Averrhoa carambola* L.

Número	Modelo	Equação	R ²
1	Primeiro grau	AFE = -10,31479 + 4,31421(C)	0,9440
2	Primeiro grau	AFE = -21,1276 + 11,4901(L)	0,9142
3	Primeiro grau	AFE = 1,435040 + 0,701533(CL)	0,9897
4	Segundo grau	AFE = 1,11661 + 0,86377(C) + 0,23962(C) ²	0,9622
5	Segundo grau	AFE = -10,4016 + 5,4819(L) + 0,8050(L) ²	0,9207
6	Segundo grau	AFE = -0,2712602 + 0,8324138(CL) - 0,0019955(CL) ²	0,9929
7	Potencia	AFE = 0,8711(C) ^{1,5873}	0,9613
8	Potencia	AFE = 1,5256(L) ^{1,9858}	0,9115
9	Potencia	AFE = 0,9911 (CL) ^{0,9203}	0,9916

Tabela 3. Coeficiente linear ($\hat{\beta}_0$), coeficiente angular ($\hat{\beta}_1$) e coeficiente de determinação (R²), obtidos das equações com ajuste linear de primeiro grau entre a área foliar estimada (AFE) pelo comprimento (C), largura (L) e produto do comprimento com a largura (CL) em função da área foliar observada (AFO), além do erro absoluto do erro (EAM), raiz quadrada médio do erro (RQME) e índice d Willmott de folhas de *Averrhoa carambola* L. utilizadas para validação

Número	$\hat{\beta}_0^{(1)}$	$\hat{\beta}_1^{(2)}$	R ²	EAM	RQME	d
1	0,7997 ^{ns}	0,9803 ^{ns}	0,9572	1,5284	1,7893	0,9882
2	-1,27948 ^{ns}	1,00985 ^{ns}	0,9271	2,2047	2,6033	0,9760
3	0,71744 [*]	0,95605 [*]	0,9940	0,5565	0,7124	0,9980
4	1,48242 [*]	0,94427 [*]	0,9698	1,3439	1,5691	0,9905
5	-0,14148 ^{ns}	0,95444 ^{ns}	0,9439	1,7816	2,1498	0,9826
6	0,08734 ^{ns}	0,98940 ^{ns}	0,9946	0,4487	0,6171	0,9985
7	1,0418 [*]	0,9660 ^{ns}	0,9700	1,2711	1,5171	0,9913
8	1,57349 [*]	0,87401 [*]	0,9520	1,5424	1,9816	0,9838
9	0,57297 [*]	0,96501 [*]	0,9948	0,4617	0,6479	0,9984

⁽¹⁾Coeficiente linear ($\hat{\beta}_0$) não difere de zero (^{ns}) e significativamente difere de zero (^{*}) pelo teste *t* de Student, ao nível de 5% de significância

⁽²⁾Coeficiente angular ($\hat{\beta}_1$) não difere de um (^{ns}) e significativamente difere de um (^{*}) pelo teste *t* de Student, ao nível de 5% de significância

Além disso, o uso de folhas pequenas, médias e grandes no ajuste da modelagem permite o uso da equação durante todo o desenvolvimento da cultura (Pezzini *et al.*, 2018), já que a área foliar de uma planta é altamente influenciada pelo tamanho das folhas, podendo ocorrer variações durante o seu ciclo (Monteiro *et al.*, 2005). Em relação à amostra utilizada para a validação, é de grande importância que os valores estejam dentro do intervalo estabelecido para a modelagem, uma vez que os dados não devem extrapolar os valores utilizados no ajuste (Levine *et al.*, 2017). Os nove modelos gerados para estimar a área foliar de carambola apresentaram significância estatística, com R² superior a 0,90 (Tabela 2). Entretanto, os modelos que utilizaram o CL apresentaram maior R² em comparação aos modelos obtidos com apenas uma das dimensões foliares (C ou L). Porém, entre todos os modelos, a equação linear de segundo grau gerada a partir de CL obteve maior significância com valor de R² (0,9929) superior aos demais modelos indicando que, a área foliar preditiva por este modelo em 99,29% dos casos é explicada por CL. A validação (Tabela 3) indica que dos nove modelos ajustados, apenas as equações 1, 2, 5 e 6, se mostraram adequadas para estimar a área de carambola, atendendo ao

critério dos coeficientes linear ($\hat{\beta}_0$) estatisticamente igual a zero e coeficiente angular ($\hat{\beta}_1$) estatisticamente igual a um. Quando o modelo segue esse comportamento, pode-se afirmar que quando a área foliar observada for zero a área estimada também será similar a zero e quando a área foliar observada tiver aumento em 1 cm², acontecerá o incremento de 1 cm² na área estimada pelo modelo (Toebe *et al.*, 2019). Os resultados também demonstram, que dentre essas quatro equações, a de modelo linear de segundo grau obtida através do produto da multiplicação do comprimento com a largura foi a que melhor estimou a área das folhas de carambola pois os valores de EAM e RQME se apresentaram mais próximos de zero e o valor do índice d de Willmott (Willmott, 1981), teve maior proximidade com um indicando maior exatidão na estimativa da área foliar. Esse modelo também apresentou melhor desempenho com valores de R² superior tanto na validação, quanto no ajuste da modelagem. Apesar da constância dos ajustes envolvendo a combinação de duas medidas da superfície foliar (Oliveira *et al.*, 2017; Schmildt *et al.*, 2017), em campo, a obtenção de duas medidas se torna mais exaustiva, exigindo maior tomada de tempo na leitura dos

dados, entretanto, esses modelos são mais precisos, pois o comprimento e largura são medidas discrepantes, assim, sua combinação se torna mais adequada do que quando mesuradas individualmente (Lavanhole et al., 2018). Diante disso, verificou-se que a equação de modelo linear de segundo grau.

$AFE = -0,2712602 + 0,8324138(CL) - 0,0019955(CL)^2$ cujo comportamento esta na Fig. 2, melhor prediz a área das folhas de carambola. Esse modelo foi gerado através do produto da multiplicação do comprimento com a largura das folhas. Além disso, é possível observar através do diagrama de dispersão do resíduo que mais de 99% dos dados se encontram entre ± 3 desvios padrões que indica acurácia na estimativa da área foliar por parte desta equação, já que quando ela for usada apenas 1% das folhas, devido ao acaso, podem apresentar valores errôneos da área foliar (Almeira, 2019). Deve-se ressaltar, que essa equação atendeu a todos os critérios estatísticos avaliados neste estudo podendo estimar que maneira precisão a área das folhas de carambola, auxiliando em futuros estudos que se deseja obter respostas a estímulos, como a adoção de fertilizantes, intensidade luminosa e a disponibilidade hídrica. Assim, apesar do método de digitalização de imagens, como utilizados neste estudo apresentar alta acurácia e baixo custo (Bosco et al., 2012), ainda há necessidade da destruição das folhas. Entretanto, após o ajustar das equações, pesquisas subsequentes podem ser feitas sem nova destruição, já que as medidas do comprimento e da largura podem ser obtidas de maneira simples, não sendo necessário o uso de equipamentos exclusivos para tal atividade.

Conclusão

O modelo linear de segundo grau foi o mais adequado para estimar a área de carambola, sendo representado pela equação $AFE = -0,2712602 + 0,8324138(CL) - 0,0019955(CL)^2$, obtida a partir do produto da multiplicação do comprimento com a largura.

REFERÊNCIAS

- Almeida, C. O. 2019. Ensaios: da amostra ao teorema do limite central um pouco dos fundamentos e uma aplicação prática. *Documentos/ Embrapa Mandioca e Fruticultura*, pp. 40.
- Alvares, C. A.; Stape, J. L.; Sentelhas, P. C.; Gonçalves, J. L. M.; Sparovek, G. 2014. Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, 22, pp. 711-728.
- Andrade, R. A.; Martins, A. B. G. 2007. Aspectos morfológicos de folhas na diferenciação de variedades de carambola. *Revista Brasileira de Fruticultura*, 29, pp. 386-388.
- Astegiano, E. D.; Favaro J. C; Bouzo, C. A. 2001. Estimación del área foliar em distintos cultivares de tomate (*Lycopersicon esculentum* Mill.) utilizando medidas foliares lineales. *Investigation Agrícola*, 16, pp.245-256.
- Bastos, D. C. 2004. A cultura da carambola. *Revista Brasileira de Fruticultura*, 26, pp. 193 – 384.
- Bosco, L.C.; Bergamaschi, H.; Cardoso, L.S.; Paula, V.A.; Casamali, B. 2012. Seleção de modelos de regressão para estimar a área foliar de macieiras 'Royal Gala' e 'Fuji Suprema' sob tela antigranizo e em céu aberto. *Revista Brasileira de Fruticultura*, 34, pp.504-514.
- Buttaro, D.; Roupael, Y.; Rivera, C. M.; Colla, G.; Gonnella, M. 2015. Simple and accurate allometric model for leaf area estimation in *Vitis vinifera* L. genotypes. *Photosynthetica*, 53, 342-348.
- Core Team. R 2018. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
- Ferreira, E. B.; Cavalcanti, P. P.; Nogueira, D. A. 2018. Package 'ExpDes.pt'.
- Gonzalez-Sanpedro, M. C.; Toan, T. le; Moreno, J.; Kergoat, L.; Rubio, E. 2008 Seasonal variation of leaf area index of agricultural field retrieved from Landsat data. *Remote Sensing of Environment*, 112, pp.810-824.
- Lavanhole, D. F.; Oliveira, P. S.; Vitória, E. L.; Aoyama, E. M. 2018. Estimativa de área foliar por meio de relações alométricas em *Aechmea blanchetiana* (Baker) L. B. SM sob distintas condições de luminosidade. *Iheringia: Série Botânica*, 73, pp. 363-373.
- Levine, D. M., Stephan, D. F.; Szabat, K. A. 2017. Estatistic for managers using Microsoft Excel: global edition. (8th ed.) London: Person. pp. 1-728.
- Monteiro, J. E. B. A.; Sentelhas, P. C.; Chiavegato, E. J.; Guiselini, C.; Santiago, A. V.; Prella, A. 2005. Estimación da área foliar do algodoeiro por meio de dimensões e massa das folhas. *Bragantia*, 64, pp.15-24.
- Oliveira, P. S.; Silva, W.; Costa, A. A. M.; Schmidt, E. R.; Vitória, E. L. 2017. Leaf area estimation in litchi by means of allometric relationships. *Revista Brasileira de Fruticultura*, 39, pp. 1-6.
- Oliveira, V. S.; Hell, L. R.; Santos, K. T. H.; Pelegrini, H. R.; Santos, J. S. H.; Oliveira, G. E.; Nascimento, A. L.; Santos, G. P.; Schmidt, O.; Czepak, M. P.; Arantes, S. D.; Alexandre, R. S.; Schmidt, E. R. (2019). Estimation of Leaf Area of Jackfruit Through Non-destructive Method. *Journal of Agricultural Science*, 11, pp.77- 85.
- Pezzini, R. V.; Cargnelutti Filho, A.; Alves, B. M.; Follmann, D. N.; Kleinpaul, J. A.; Wartha, C. A.; Silveira, D. L. (2018) Models for leaf area estimation in dwarf pigeon pea by leaf dimensions. *Bragantia*, 77, pp. 221-229.
- Reis, L., S.; Azevedo, C. A. V.; Albuquerque, A. W.; Silva Júnior, J. F. 2013. Índice de área foliar e produtividade do tomate sob condições de ambiente protegido. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 17, pp.386–391.
- Schindelin, J.; Rueden, C. T.; Hiner, Mc.; Eliceiri, K. W. 2015. The ImageJ Ecosystem: An Open Platform for Biomedical Image Analysis. *Molecular Reproduction and Development*. 82, pp 518–529.
- Schmidt, E. R.; Hueso, J. J.; Pinillos, V.; Stedfeldt, A.; Cuevas, J. 2017. Allometric models for determining leaf area of 'Fino de Jete' chirimoy grown in greenhouse and in the open field. *Fruits*, 72, pp. 24-30.
- Silva, C. A. 2016. Correlações fenotípicas e análise de trilha em caracteres morfoagronômicos de mamoeiro. *Revista Agro@ambiente On-line*, 10, pp. 217-227.
- Toebe, M.; Souza, R.R.; Mello, A.C.; Melo, P.J.; Segatto, A.; Castanha, A.C. (2019) Leaf area estimation of squash 'Brasileirinha' by leaf dimensions. *Ciência Rural*, 49, pp. 1-11
- Vitória, E. L.; Freitas, I. L. J, Locatelli, T., Lacerda, E. G., Valle, J. M., Pereira, R. C., Almeida, P. F. P., Vitoria, R. Z., Simom, C. P, Fernandes, A. A. 2018. Mathematical Models for Leaf Area Estimates of Guava. *Journal of Agricultural Science*, 10, pp.272-278.
- Willmott, C. J. 1981. On the validation of models. *Physical Geography*, 2, pp. 184-194.