

## Estimativa da área foliar do antúrio com o uso de funções de regressão

### Estimating leaf area in anthurium with regression functions

Silvia Helena Modenese-Gorla da Silva<sup>1\*</sup> Juliana Domingues Lima<sup>1</sup>  
Hugo do Nascimento Bendini<sup>1</sup> Edson Shigueaki Nomura<sup>11</sup>  
Wilson da Silva Moraes<sup>11</sup>

- NOTA -

#### RESUMO

*O presente trabalho teve como objetivo determinar quais variáveis dimensionais da folha são mais adequadas para utilização na estimativa da área foliar do antúrio (*Anthurium andraeanum*), cv. "Apalai", por meio de equação de regressão linear, e comparar o desempenho de diferentes funções de regressão obtidas com o uso de aprendizado de máquina (AM). A variável que melhor estimou a área foliar foi o produto das dimensões lineares (comprimento e largura), CxL, sendo a equação proposta  $A_f = 0.9672 * C \times L$ , com coeficiente de determinação ( $R^2$ ) de 0,99. Verificou-se, também, com o uso de AM, que as funções lineares são mais adequadas para a estimação da área foliar dessa espécie vegetal.*

**Palavras-chave:** plantas ornamentais, análise de crescimento, dimensões foliares.

#### ABSTRACT

*This study was aimed at determining which of the leaf dimensional parameters, length (C), width (L) or the product of the length by width (CxL), are more appropriate to estimate area of anthurium (*Anthurium andraeanum*), cv. Apalai using a linear regression equation and to compare the performance of different regression functions obtained with machine learning (AM). The parameter that estimated the leaf area with the least error was the CxL product, with the equation proposed  $A_f = 0.9672 * C \times L$  and coefficient of determination ( $R^2$ ) equal to 0.99. It was also verified with AM that linear functions are more adequate to estimate the leaf area of this vegetal species.*

**Key words:** ornamental plants, growth analysis, leaf dimensions.

No Brasil, o antúrio (*Anthurium andraeanum*) tem se destacado como espécie importante para produção de folhas e flores de corte e para cultivo em vaso. Devido ao importante papel das plantas ornamentais na diversificação da agricultura tropical, existe grande interesse em conhecer características do crescimento e desenvolvimento dessa espécie vegetal visando a melhorar seu potencial produtivo.

A área foliar é um dos principais parâmetros do crescimento vegetal, pois está relacionada com diversos processos fisiológicos da planta, tais como fotossíntese, respiração e transpiração. Dentre os métodos conhecidos de determinação da área foliar, merece destaque o procedimento que relaciona a área foliar com as dimensões lineares da folha (PEREIRA, 1987), devido à rapidez na obtenção de dados, quando se utiliza amostra de folhas, e por não ser, necessariamente, destrutivo (PEDRO-JÚNIOR et al., 1986).

No intuito de obterem-se dados de área foliar confiáveis, faz-se necessário a determinação de qual das dimensões foliares, comprimento (C), largura (L) ou seu produto (CxL), é mais adequada para estimativa da área foliar real da planta. Para isso, são ajustadas funções de regressão baseadas nas dimensões lineares da folha.

A escolha da função de regressão que melhor estima a área foliar, geralmente, é baseada no

<sup>1</sup>Campus Experimental de Registro, Universidade Estadual Paulista "Júlio de Mesquita Filho" (UNESP). Rua Tamekishi Takano, nº 5, Centro, 119000-000, Registro, SP, Brasil. E-mail: silvia@registro.unesp.br. \*Autor para correspondência.

<sup>11</sup>Pólo Regional de Desenvolvimento Sustentável dos Agronegócios do Vale do Ribeira, Agência Paulista de Tecnologia dos Agronegócios (APTA), Pariquera-Açu, SP, Brasil.

coeficiente de determinação ( $R^2$ ) ou erro padrão. Contudo, podem-se utilizar ferramentas alternativas tais como técnicas de Inteligência Artificial (IA) (REZENDE, 2003; MITCHELL, 1997).

O Aprendizado de Máquina (AM) é uma área da IA cujo objetivo é o desenvolvimento de técnicas computacionais sobre o aprendizado, bem como a construção de sistemas capazes de adquirir conhecimento de forma automática (MONARD & BARANAUSKAS, 2003). Assim, as técnicas de AM, por serem capazes de aprender por si próprias (MITCHELL, 1997), também representam uma alternativa atraente para lidar com problemas de ajuste de funções, ou regressões. Diversos trabalhos aplicaram o AM na área da agronomia (MICHALSKI & CHILAUSSKY, 1980; McQUEEN et al., 1995; MODENESE-GORLADASILVA, 2005).

Assim, o trabalho teve como objetivo determinar quais das variáveis dimensionais, comprimento (C), largura (L), produto do comprimento pela largura (CxL), são as mais adequadas para utilização na estimativa da área foliar do antúrio por meio de equação de regressão linear, e, posteriormente, comparar o desempenho de diferentes funções de regressão obtidas com o uso de AM.

Para tal, foram coletadas aleatoriamente 200 folhas de tamanhos variados de plantas de *Anthurium andraeanum*, cultivar "Apalai", que se encontravam em área de produção comercial em Pariqueira-Açu, SP. No laboratório do Campus Experimental de Registro, Universidade Estadual Paulista, foi determinado o comprimento máximo (C) e a largura máxima (L), perpendicular à nervura principal do limbo foliar, e a área foliar (Af) com o medidor portátil de área foliar Licor Mod. LI - 3000.

A partir do conjunto de dados obtidos, foram ajustados modelos lineares para estimar a área do limbo (Af) do tipo  $Y = bX$ , em que o valor Y corresponde à Af estimada em função das dimensões da folha, comprimento (C), largura (L) e produto (CxL). A base de dados foi composta por 200 padrões e quatro atributos contínuos. Para estudo exploratório da base de dados e para determinação da função de regressão, foi utilizada a linguagem "R" (RDEVELOPMENT CORE TEAM, 2005). A comparação das equações obtidas com as variáveis dimensionais foi feita por meio dos valores dos coeficientes de determinação da equação ( $R^2$ ) e do erro padrão.

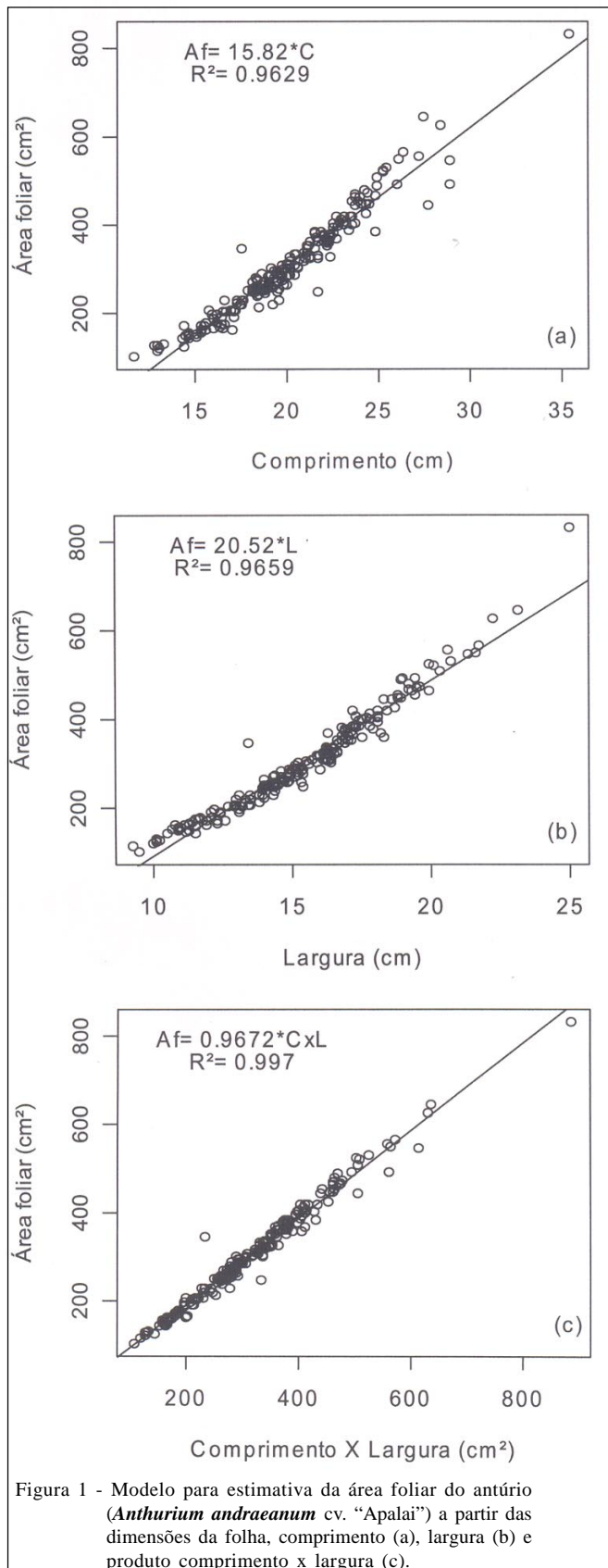
O estudo de AM foi realizado por meio da ferramenta computacional WEKA. Os resultados foram obtidos utilizando *10-fold cross validation* (MITCHELL, 1997). O método de amostragem *10-fold cross validation* particiona o conjunto de dados total

em dez partes iguais, utilizando nove partes para treinar e uma para testar. Foram realizados, portanto, dez diferentes treinamentos e testes, repetindo-se cada *10-fold cross validation* dez vezes, totalizando, então, 100 treinamentos para cada algoritmo testado. Para o treinamento, foram utilizadas Redes Neurais Artificiais (RNAs) *Perceptron* multicamadas, treinadas com o algoritmo *backpropagation* (RUMELHART & McCLELLAND, 1986; BRAGA et al., 2000), com termo *momentum*. A taxa de aprendizado e o termo *momentum* adotados foram respectivamente iguais a 0 e 1. Também foram utilizadas RNAs com uma única camada, treinadas com o algoritmo mínimo quadrado médio (*LMS – Least mean square*) (HAYKIN, 1999). Para comparação do estudo da função de regressão linear com as de AM, utilizou-se a raiz quadrada do erro quadrático médio (*root mean square error*) obtida nos conjuntos de teste.

A análise de frequência dos dados de área foliar obtida com o uso do medidor mostrou que metade dos dados está mais concentrada entre 100 e 290cm<sup>2</sup> e a outra metade encontra-se mais dispersa entre 290 e 800cm<sup>2</sup>. Os valores do comprimento (C) das folhas variaram de 11,70 a 35,50cm, com valor médio de 19,98cm, e a largura máxima das folhas variou de 9,30 a 25,00cm, com valor médio de 15,41cm. Para a área foliar obtida com o medidor, os valores variaram de 103,7 a 831,4cm<sup>2</sup>, com média de 306,2cm<sup>2</sup>.

Os modelos de regressão linear obtidos com a reta ajustada para as variáveis C, L e CxL foram apresentados na figura 1. Os erros-padrão foram de 63,18, 60,55 e 17,96cm<sup>2</sup>, respectivamente, para os modelos obtidos a partir das variáveis C, L e CxL. Esses resultados mostram que os modelos para estimativas da área foliar foram adequados, considerando os altos coeficientes de determinação ( $R^2$ ), próximos de 1, e os baixos valores de erro padrão apresentados pelos mesmos. A equação linear obtida a partir da variável CxL,  $Af = 0,9672 * CxL$ , foi a que apresentou menor erro padrão e maior  $R^2$ , indicando que o produto CxL e o modelo de regressão linear passando pela origem são indicados para estimativa da área foliar da cultivar de antúrio estudada. O  $R^2$  de 0,99 indica que o modelo de regressão linear pela origem estimou a área foliar com uma precisão de 99%.

Em outras espécies vegetais como *Typha latifolia* (BIANCO et al., 2002), morangueiro (STRIK & PROCTOR, 1985; PIRES et al., 1999) e em variedades do gênero *Castanea sativa* (SERDAR & DEMIRSOY, 2006), a utilização em conjunto do comprimento e da largura (CxL) também mostrou os melhores resultados. No entanto, para uva cultivares "Niágara" e "DeChaunac" (LIVY WILLIAMS III & MARTINSON,



2003) e para meloeiro (NASCIMENTO et al., 2002), apenas o uso da largura mostrou ser mais adequada para estimar a área foliar.

Na seleção do atributo de entrada a ser utilizado (CxL) para determinação da área foliar com o uso de função de regressão linear, observou-se que as variáveis C e L apresentavam uma discreta relação não-linear com a área foliar (Figura 1). Assim, na comparação da função de regressão linear com os algoritmos de AM, optou-se por testar também as variáveis C e L, de maneira a verificar o desempenho não-linear com o uso do algoritmo *backpropagation*. As funções de Regressão *LMS*, *backpropagation* e linear simples apresentaram raiz quadrada do erro quadrático médio (RMSE) e desvio padrão, para os conjuntos de teste da variável CxL, respectivamente iguais a  $16,57 \pm 6,66$ ;  $19,09 \pm 7,46$  e  $16,56 \pm 6,34$ . Para a variável C, foram obtidos valores respectivamente iguais a  $28,76 \pm 6,96$ ;  $33,45 \pm 12,97$  e  $28,48 \pm 7,17$ , enquanto que, para a variável L, iguais a  $24,36 \pm 7,69$ ;  $23,51 \pm 12,58$  e  $23,67 \pm 7,50$ .

Os menores erros e desvios-padrão foram apresentados para a variável CxL, com as funções lineares. Pode-se dizer, com 95% de confiança, que a variável CxL foi a que melhor estimou a área foliar do antúrio. Também se verifica, com 95% de confiança, que, para as variáveis C e CxL, as funções lineares apresentaram melhor desempenho quando comparadas com o algoritmo *backpropagation*, demonstrando que a estimativa da área foliar constitui um problema com resolução linear. Também, comprovou-se que o algoritmo *backpropagation* obteve melhor desempenho na resolução de problemas de funções não-lineares, conforme reportado por HAYKIN (1999). Dessa maneira, de acordo com a avaliação dos modelos de regressão utilizada (RMSE), a equação linear passando pela origem determinada apresenta uma boa estimativa da área foliar do antúrio.

Os resultados encontrados no presente trabalho permitem concluir que a variável que melhor estima a área foliar de *Anthurium andraeanum* cv. "Apalai" é o produto do comprimento pela largura da folha (CxL), a partir da equação  $Af = 0,9672 * C * L$ . O estudo com aprendizagem de máquina mostrou que as funções lineares são mais adequadas para a estimativa da área foliar

dessa espécie vegetal, apresentando-se como uma ferramenta alternativa na busca de melhores funções de regressão.

## AGRADECIMENTOS

À Fundação para o Desenvolvimento da Universidade Estadual Paulista (FUNDUNESP), pelo apoio financeiro.

## REFERÊNCIAS

- BIANCO, S. et al. Estimativa da área foliar de *Typha latifolia* usando dimensões lineares do limbo foliar. **Planta Daninha**, v.20, n.3, p.353-356, 2002.
- BRAGA, A.P. et al. **Redes neurais artificiais: teoria e aplicações**. Rio de Janeiro: Livro Técnico e Científico, 2000. 262p.
- HAYKIN, S. **Neural networks: a comprehensive foundation**. 2.ed. Hamilton: Prentice Hall, 1999. 842p.
- LIVY WILLIAMS, III; MARTINSON, T.E. Nondestructive leaf area estimation of 'Niagara' and 'DeChaunac' grapevines. **Scientia Horticulturae**, v.98, n.4, p.493-498, 2003.
- McQUEEN, R.J. et al. Applying machine learning to agricultural data. **Computers and Electronics in Agriculture**, v.12, p.249-261, 1995.
- MICHALSKI, R.S.; CHILAUSSKY, R.L. Learning by being and learning from examples: an experimental comparison of the two methods of knowledge acquisition in the context of developing an expert system for soybean disease diagnosis. **International Journal of Policy Analysis Information Systems**, v.4, n.2, p.125-161, 1980.
- MITCHELL, T.M.; **Machine learning**. New York: McGraw-Hill, 1997. 414p.
- MODENESE-GORLA DA SILVA, S.H. et al. Estudo da Influência de fatores ambientais em frangos de corte com Redes Neurais Artificiais. In: SIMPÓSIO BRASILEIRO DE AUTOMAÇÃO INTELIGENTE, 7.; E IEEE LATIN AMERICAN ROBOTICS SYMPOSIUM, 2., 2005, São Luís do Maranhão, MA. **Anais...** São Luís do Maranhão: UFMA, 2005. 1 CD.
- MONARD, M.C.; BARANAUSKAS, J.A. Conceitos de aprendizado de máquina. In: REZENDE, S.O. **Sistemas inteligentes: fundamentos e aplicações**. Barueri, SP: Manole, 2003. Cap.4, p.89-114.
- NASCIMENTO, I.B. do et al. Estimativa da área foliar do meloeiro. **Horticultura Brasileira**, Brasília, v.20, n.4, p.555-558, 2002.
- PEDRO-JÚNIOR, M.J. et al. Determinação da área foliar em videira cultivar Niágara Rosada. **Bragantia**, Campinas, v.45, n.1, p.199-204, 1986.
- PEREIRA, A.R. Estimativa da área foliar em milharal. **Bragantia**, Campinas, v. 46, n.1, p.147-150, 1987.
- PIRES, R.C.M. et al. Estimativa da área foliar do morangueiro. **Horticultura Brasileira**, Brasília, v.17, n.2, p.86-90, 1999.
- R DEVELOPMENT CORE TEAM. **R: a language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2005. 2584p.
- REZENDE, S.O. Introdução. In: REZENDE, S.O. **Sistemas inteligentes: fundamentos e aplicações**. Barueri, SP: Manole, 2003. Cap. 1, p. 3-8.
- RUMELHART, D. E., MCCLELLAND, J. L. **Parallel distributed processing: foundations**. Cambridge: MIT, 1986. V.1, 547p.
- SEDAR, Ü; DEMIRSOY, H. Non-destructive leaf area estimation in chestnut. **Scientia Horticulturae**, Amsterdam, v.108, n.2, p.227-230, 2006.
- SPIEGEL, M.R. **Estatística**. 3.ed. São Paulo: Pearson Education do Brasil, 1994. 643p.
- STRIK, B.C.; PROCTOR, J.T.A. Estimating the area of trifoliate and unequally imparipinnate leaves of strawberry. **HortScience**, Alexandria, v.20, n.6, p.1072-1074, 1985.