

Método de comparação de modelos de regressão não-lineares em bananeiras

Method of comparison of models non-linear regression in bananas trees

Emanuel Maia^I Dalmo Lopes de Siqueira^{II} Fabyano Fonseca e Silva^{III}
Luiz Alexandre Peternelli^{III} Luiz Carlos Chamhum Salomão^{II}

RESUMO

A utilização de modelos matemáticos para descrever o crescimento de vegetais possui uma larga aplicação na pesquisa agropecuária. Nesse tipo de pesquisa, dois problemas são frequentes: encontrar o modelo que melhor se ajusta ao fenômeno biológico e identificar as curvas entre os tratamentos que sejam estatisticamente semelhantes. Assim, o presente trabalho teve por objetivos comparar a qualidade do ajuste de modelos de regressão não-linear e comparar as curvas provenientes desse melhor modelo em relação aos diferentes tratamentos. Para essa análise, foram utilizados os dados de um experimento instalado no arranjo fatorial 2x5, sob o delineamento em blocos casualizados. Os níveis dos fatores foram constituídos por duas cultivares de bananeira e cinco doses de um regulador de crescimento. A comparação das curvas de diferentes tratamentos foi realizada a partir da análise de agrupamento com o auxílio de três estatísticas que permitiam acessar o número ideal de grupos. Na avaliação das curvas de crescimento, o modelo logístico demonstrou-se superior. Na comparação dos tratamentos, verificou-se que o método da análise de agrupamento, com as modificações propostas, possibilitou detectar três grupos de tratamentos, assemelhando-se ao comportamento biológico.

Palavras-chave: *Musa spp.*, análise de crescimento, modelos não lineares, análise de agrupamentos, modelo logístico.

ABSTRACT

The use of models to describe the growths of plants has a wide application in the agronomy and cattle research. In this type of research, two problems are often present: to find the model that shows more information, which is best fitted to the

biologic phenomenon and identifying curves that are statistically equal. Thus, the objective of the present work was to compare the quality adjustments models of non linear regression in the description of growth and to compare the curves from this best model in relation to different treatments. Data from an experiment installed in a factorial arrangement 2x5, in a randomized design of blocks with eight repetitions were used. The levels of treatments were constituted by two banana tree cultivars and five doses of a growth regulator. The cluster analysis was used with three statistics that allowed access to the ideal number of groups. In the comparison of growth models, the superiority of the logistic model was observed. For comparison of treatments, it was verified that the cluster method analysis with the changes proposed allowed to detect three groups of treatments, that are similar to the biologic behavior.

Key words: *Musa spp.*, growth analysis, non-linear models, cluster analysis, logistic model.

INTRODUÇÃO

O estudo do crescimento de vegetais possui uma larga aplicação na pesquisa agropecuária, pois possibilita ao pesquisador comparar o comportamento de uma planta ou de suas partes de interesse, diante das diversas situações experimentais empregadas. Para esse estudo, têm sido utilizados vários modelos lineares e não-lineares para descrever crescimento ao longo do tempo (ALVAREZ & CASTRO, 1999; URCHEI et al., 2000; HERNÁNDEZ et al., 2007). Contudo, observa-se que os modelos não-lineares proporcionam vantagens

^IDepartamento de Agronomia, Universidade Federal de Rondônia (UNIR), 76940-000, Rolim de Moura, RO, Brasil. E-mail: emanuelfms@gmail.com. Autor para correspondência.

^{II}Departamento de Fitotecnia, Setor de Fruticultura, Universidade Federal de Viçosa (UFV), Viçosa, MG, Brasil.

^{III}Departamento de Informática, Setor de Estatística, UFV, Viçosa, MG, Brasil.

em relação aos modelos lineares, os quais podem auxiliar o pesquisador no processo de tomada de decisão por apresentarem parâmetros com interpretação biológica. Dentre essas funções não-lineares, pode-se citar os modelos de: Brody ou Monomolecular (BRODY, 1945), Gompertz (WINSOR, 1932), Logístico (NELDER, 1961) e von Bertalanffy (von BERTALANFFY, 1957). Assim, a tarefa de encontrar o modelo que melhor se ajuste ao fenômeno é relatada na literatura como árdua (BARROS et al., 1984; ENGEL & POGGIANI, 1990; OLIVEIRA et al., 2000; SANTORO et al., 2005).

Em muitas situações experimentais, também existe o interesse em verificar se duas ou mais curvas provenientes de diferentes tratamentos possuem características estatisticamente similares. Há, na literatura, propostas para a comparação de curvas de regressão não-lineares (REGAZZI, 2003; REGAZZI & SILVA, 2004). Contudo, para conjuntos de dados nos quais há muitos tratamentos ou os dados são provenientes de delineamentos complexos, essas propostas são limitadas. Dessa forma, supõe-se que a análise de agrupamento, utilizando os parâmetros das regressões como variáveis dependentes, poderia ser utilizada para determinar quais tratamentos possuem curvas similares.

Este trabalho teve por objetivos comparar a qualidade do ajuste de modelos de regressão não-linear na descrição do crescimento e comparar as curvas provenientes do melhor modelo em relação a duas variedades de bananeiras submetidas a cinco doses de um regulador de crescimento.

MATERIAL E MÉTODOS

Os dados são provenientes de um experimento para avaliar o efeito do paclobutrazol, regulador de crescimento, sobre o desenvolvimento de bananeiras (SOUZA, 2007). Este foi instalado de acordo com o arranjo fatorial 2x5, sob o delineamento em blocos casualizados, com oito repetições. Os níveis dos fatores foram: duas cultivares de bananeira ('Prata Anã' e 'FHIA 01') e cinco doses de paclobutrazol (0,0; 0,5; 1,0; 1,5 e 2,0g de i.a. planta⁻¹) aplicado via solo. A variável utilizada foi a altura do pseudocaule das bananeiras (cm), medida do solo até a segunda folha da roseta foliar, compreendendo o período do transplântio das mudas micropropagadas até a emissão da ráquis.

Os modelos utilizados para descrever o crescimento foram: Brody, $Y_i = a - be^{-kti}$; Gompertz, $Y_i =$

$ae^{-be^{-kti}}$; Logístico, $Y_i = a(1 + e^{-kti})^{-1}$; von Bertalanffy, $Y_i = a(1 - be^{-kti})^3$, em que: Y_i representa a altura do pseudocaule observado no tempo t_i ; t_i representa o tempo decorrido em meses do transplântio até a aferição; a representa a altura assintótica; b representa o parâmetro de escala; e k representa a taxa média de crescimento (BRODY, 1945; WINSOR, 1932; von BERTALANFFY, 1957; NELDER, 1961).

Foi ajustada uma curva de crescimento para cada repetição, totalizando 80 curvas, pelo PROC MODEL do software estatístico SAS® (SAS INSTITUTE, 2003). Para a escolha do modelo mais adequado, foram utilizados os seguintes avaliadores de qualidade de ajuste: quadrado médio do resíduo (QME), porcentagem de convergência (%CG), *Bayesian Information Criterion* (BIC) e coeficiente de determinação (R^2 , em que $R^2 = \text{SQRegressão}/\text{SQTotal}$), comparados pelo intervalo de confiança ao nível de 95% de probabilidade.

Os parâmetros obtidos pelo modelo mais adequado pelos critérios adotados foram considerados como variáveis dependentes em uma análise de variância multivariada, PROC GLM (SAS INSTITUTE, 2003), com a opção MANOVA, com o intuito de avaliar o efeito dos tratamentos sobre as estimativas dos parâmetros (a , b e k). Sendo constatado esse efeito pela significância do teste do método do traço de Pilai, extraíram-se as médias ajustadas pelo método dos quadrados mínimos (LSMEANS). Estas foram submetidas à análise de agrupamento PROC CLUSTER (SAS INSTITUTE, 2003), utilizando método de Ward por utilizar como critério para a construção do dendograma a soma de quadrados entre os grupos formados (KHATTREE & NAIK, 2000).

Foram utilizadas três estatísticas que permitem acessar o número ótimo de grupos: a *Root Mean Square* (RSQ), que é a razão entre a soma de quadrados entre grupos e a soma de quadrados total, e mede a heterogeneidade do agrupamento; a *Semi-Partial Root Mean Square* (SPRSQ), que mede a perda de homogeneidade por juntar dois grupos; e a *Between-group Sum of Squares* (BSS), que é a soma de quadrados entre grupos para dois agrupamentos, e mede a homogeneidade de grupos unidos (KHATTREE & NAIK, 2000). Com esses valores (RSQ, SPRSQ e BSS) como variáveis dependentes e o número de grupos formados em cada estágio da análise de

agrupamento, como variável regressora, ajustou-se um modelo de regressão segmentado Quadrático Response Platô. Esse modelo assume que, antes de um ponto ótimo, a variável dependente (nesse caso, as estatísticas RSQ, SPRSQ e BSS) cresce ou decresce em função da variável independente (nesse caso, número de grupos) e depois desse ponto ótimo a função mantém respostas constantes. No presente trabalho, esse ponto ótimo representa o número de grupos a partir do qual as estatísticas mencionadas não se alteram mais, ou seja, mesmo aumentando o número de cluster, a precisão é a mesma. Mais detalhes sobre o modelo de regressão segmentado Quadrático Response Platô podem ser obtidos em REZENDE et al. (2007).

O ponto de interseção entre os seguimentos do modelo, ou ponto ótimo (X_0), identifica o número ótimo de grupos em relação à estabilização dos valores das estatísticas, visto que o número apropriado de grupos seria aquele em que as mudanças nesses valores são mínimas, exceto para a RSQ, em que são permitidas mudanças com valores mais elevados (KHATTREE & NAIK, 2000).

RESULTADOS E DISCUSSÃO

Todos os modelos estudados obtiveram 100% de convergência, e os demais avaliadores de qualidade de ajuste estão apresentados na tabela 1. Comparando-se dois modelos, aquele com menor valor de BIC deve ser o adotado (SCHWARZ, 1978). Os valores do BIC para o modelo logístico são, em média, 40% menores que os encontrados para o modelo de Gompertz. As diferenças na qualidade de ajuste entre o modelo logístico e os demais se acentuam para as estimativas do quadrado médio do resíduo (QME). As menores diferenças foram encontradas para o R^2 , em

que o modelo Logístico apresentou os maiores valores entre os pares. Devido à superioridade na qualidade de ajuste pelos indicadores utilizados, o modelo logístico foi adotado para dar continuidade ao objetivo do presente trabalho. Os resultados da análise de variância multivariada, utilizando as estimativas dos parâmetros do modelo logístico como variáveis dependentes, mostraram efeito significativo do efeito de tratamentos pelo método do traço de Pilai. Dessa forma, foi realizada a análise de agrupamento, cujos resultados estão apresentados na tabela 2.

Pelo modelo de regressão segmentado Quadrático Response Platô, o ponto de interseção (X_0) identifica o número ótimo de grupos, assim os resultados obtidos na análise indicaram, em geral, valores próximos a três grupos para as estatísticas empregadas (Tabela 3). SHARMA (1996) salienta que, em razão de não serem conhecidas as distribuições amostrais dessas estatísticas, é importante observar se o comportamento biológico das características em questão está bem representado. Observando as equações ajustadas para os 10 tratamentos, observa-se a formação de quatro ou cinco grupos de respostas (Figura 1). Um grupo formado pelos tratamentos que não receberam a aplicação do paclobutrazol, sem inibição do crescimento (cultivares 'Prata Anã' e 'FHIA 01', com a dose 0,0g de i.a. de paclobutrazol). Dois ou três grupos, com respostas intermediárias, e o quarto grupo, com resposta mais acentuada da inibição do crescimento das plantas (tratamentos com as maiores doses de paclobutrazol). É importante comentar que, em média, a cultivar 'FHIA 01' apresenta maior altura do pseudocaule que a cultivar 'Prata Anã' (SOUZA, 2007), o que confirma a tendência dos tratamentos seis a 10 apresentarem maiores valores da altura assintótica

Tabela 1 - Coeficiente de determinação (R^2), Quadrado Médio do Resíduo (QME), *Bayesian Information Criterion* (BIC) e respectivos intervalos de confiança, em 95% de probabilidade ($IC_{95\%}$), para diferentes modelos não-lineares ajustados para 80 curvas de crescimento do pseudocaule (cm) de bananeiras submetidas a diferentes doses de paclobutrazol.

Modelos	----- R^2 -----			-----QME-----			-----BIC-----		
	- $IC_{95\%}$	Média	$IC_{95\%}$	- $IC_{95\%}$	Média	$IC_{95\%}$	- $IC_{95\%}$	Média	$IC_{95\%}$
Logístico	0,982	0,985	0,988	34,228	40,842	47,445	22,420	25,971	29,522
von Bertalanffy	0,959	0,964	0,968	90,939	100,644	110,531	49,574	51,681	53,788
Brody	0,916	0,922	0,928	208,400	229,289	250,180	70,017	72,217	74,417
Gompertz	0,967	0,971	0,975	70,865	79,270	87,675	42,995	45,322	47,650

Tabela 2 - Tratamentos, número de grupos formados (NG) e estatísticas de auxílio na determinação do número de grupos ('Root Mean Square', RSQ; 'Semi-Partial Root Mean Square', SPRSQ; e 'Between-group Sum of Square', BSS), em cada estágio da análise de agrupamentos, utilizando o método de Ward para curvas de crescimento de bananeiras submetidas a diferentes doses de paclobutrazol.

Tratamentos ⁽¹⁾	NG	SPRSQ	RSQ	BSS
{2, 9}, 1, 3, 4, 5, 6, 7, 8, 10	9	0,0000	1,000	0,0064
{2, 9, 8}, 1, 3, 4, 5, 6, 7, 10	8	0,0000	1,000	0,1553
{1, 6}, {2, 9, 8}, 3, 4, 5, 7, 10	7	0,0002	1,000	0,5277
{3, 10}, {1, 6}, {2, 9, 8}, 4, 5, 7	6	0,0004	0,999	1,3505
{4, 5}, {3, 10}, {1, 6}, {2, 9, 8}, 7	5	0,0006	0,999	1,7765
{4, 5, 3, 10}, {1, 6}, {2, 9, 8}, 7	4	0,0211	0,978	65,852
{4, 5, 3, 10}, {1, 6}, {2, 9, 8, 7}	3	0,0569	0,921	177,85
{4, 5, 3, 10, 2, 9, 8, 7}, {1, 6}	2	0,1594	0,761	498,25
{4, 5, 3, 10, 2, 9, 8, 7, 1, 6}	1	0,7614	0,000	2379,6

⁽¹⁾ Tratamentos dentro das chaves foram unidos pela análise de agrupamentos no estágio em questão. Tratamentos de 1 a 5 referem-se à cultivar 'Prata Anã'. Os tratamentos de 6 a 10 referem-se à cultivar 'FHIA 01'. Os tratamentos 1 e 6; 2 e 7; 3 e 8; 4 e 9; 5 e 10 referem-se respectivamente às doses 0,0; 0,5; 1,0; 1,5 e 2,0g de i.a. de paclobutrazol por planta.

estimada em relação aos tratamentos do cultivar 'Prata Anã', com as mesmas concentrações de paclobutrazol.

A análise de regressão apontou que o número ideal de grupos é igual a três (X0). Com esse número, retornou-se à tabela 2, no estágio correspondente, e foram observados os tratamentos agrupados. Os três grupos formados são apresentados na tabela 4, e as equações ajustadas são ilustradas na figura 2.

Considerando os resultados obtidos, após a análise de agrupamentos, e com o objetivo de cultivar plantas com uma menor altura do pseudocaule, verifica-se que a aplicação de 1,0g de i.a. de paclobutrazol para as bananeiras 'Prata Anã' resultaria em altura final

semelhante a aplicação de 2,0g de i.a. de paclobutrazol para as bananeiras 'FHIA 01'. Isto se deve ao fato que os parâmetros de crescimento, entre eles a altura assintótica, foram considerados idênticos, visto que esses tratamentos estão presentes no mesmo grupo. Por outro lado, esses tratamentos (e os demais presentes no grupo 3) possuem a maior taxa média de crescimento em relação aos outros grupos, que representa um menor tempo entre o plantio e o crescimento máximo das plantas.

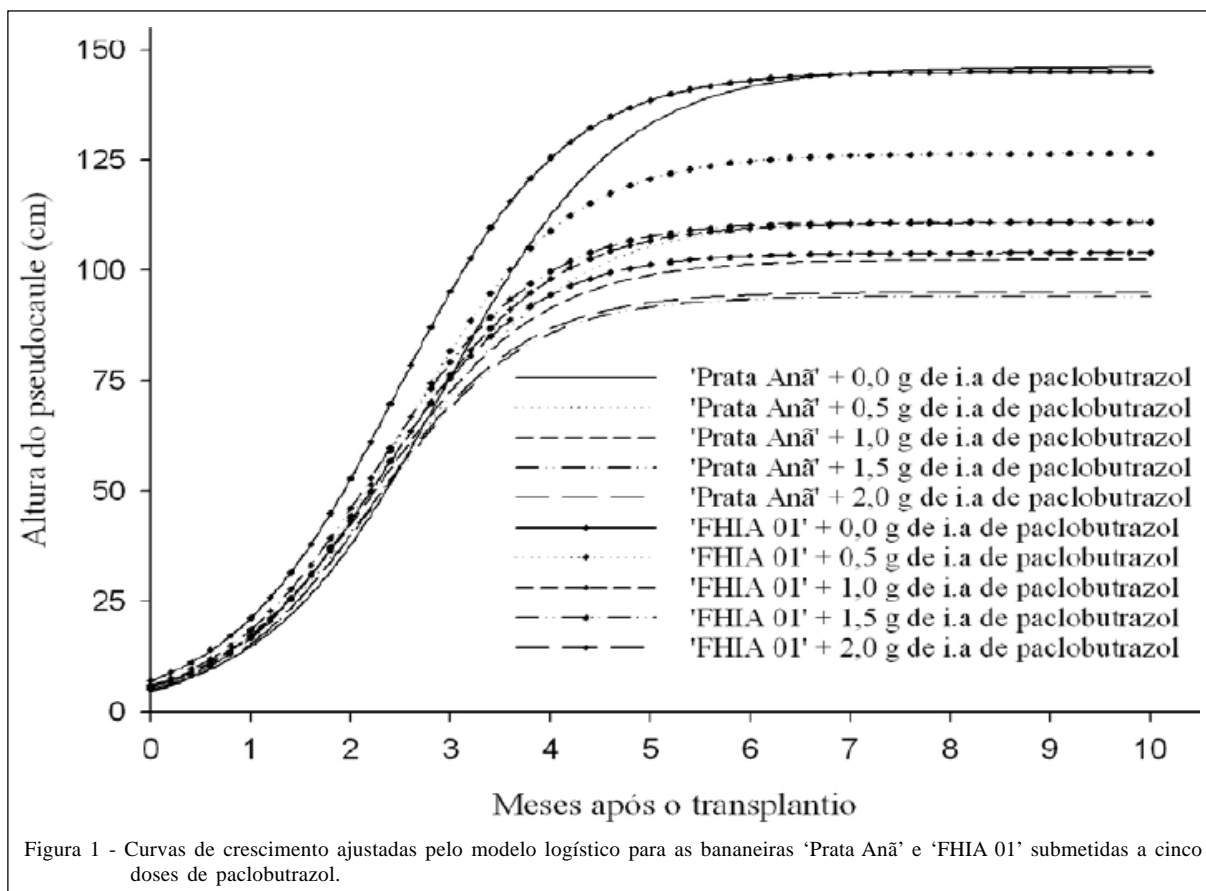
Dessa forma, verificou-se que o procedimento de utilização da análise de agrupamentos, utilizando o método de Ward, apresenta grande potencial para a comparação de curvas provenientes de diferentes tratamentos, facilitando, assim, a interpretação dos resultados pelo pesquisador. Contudo, é importante salientar que são necessários mais estudos para confirmar, em outras situações, a superioridade do modelo logístico em relação aos demais, assim como da análise de agrupamento em comparar curvas provenientes de diferentes tratamentos.

CONCLUSÕES

O modelo logístico apresentou melhor qualidade de ajuste e pode ser utilizado para descrever curvas de crescimento de bananeiras. O método de análise de agrupamento apresenta potencial na

Tabela 3 - Parâmetros estimados para a regressão segmentada (A, B, C, X0, Platô, R²), para as três estatísticas (Root Mean Square, RSQ; Semi-Partial Root Mean Square, SPRSQ e Between-group Sum of Square, BSS) utilizadas no método de Ward da análise de agrupamento de curvas de crescimento de bananeiras tratadas com diferentes doses de paclobutrazol.

Parâmetros	RSQ	SPRSQ	BSS
A	-1,291	1,818	5880
B	1,556	-1,283	-4010,2
C	-0,265	0,227	709,600
X0	2,937	2,826	2,826
Platô	0,994	0,005	14,560
R ²	0,340	0,625	0,308



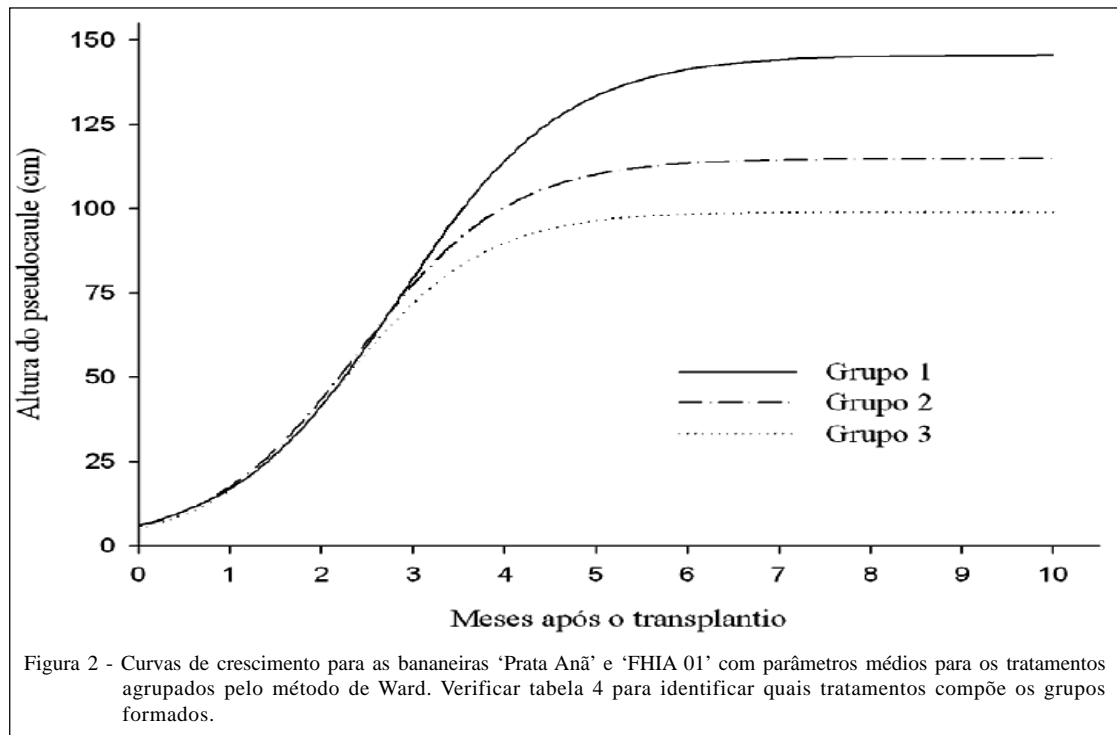
comparação das curvas de crescimento e possibilitou detectar três grupos de tratamentos semelhantes em relação à forma da curva de crescimento das bananeiras estudadas.

AGRADECIMENTO

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico, pela concessão da bolsa de estudos ao primeiro autor.

Tabela 4 - Grupos formados pela análise de agrupamento usando o método de Ward para as curvas de crescimento e os parâmetros médios estimados (a, b e k) do modelo logístico, considerando a altura do pseudocaule (cm) de bananeiras submetidas a diferentes doses de paclobutrazol.

Grupo	Tratamentos	a	b	k
1	'Prata Anã' + 0,0g de i.a. de paclobutrazol	146,013	3,305	1,133
1	'FHIA 01' + 0,0g de i.a. de paclobutrazol	145,049	2,967	1,025
Média(G1)		145,531	3,136	1,079
2	'Prata Anã' + 0,5g de i.a. de paclobutrazol	111,182	2,849	1,149
2	'FHIA 01' + 0,5g de i.a. de paclobutrazol	126,407	3,079	1,228
2	'FHIA 01' + 1,0g de i.a. de paclobutrazol	110,695	2,944	1,246
2	'FHIA 01' + 1,5g de i.a. de paclobutrazol	111,152	2,863	1,257
Média(G2)		114,859	2,934	1,220
3	'Prata Anã' + 1,0g de i.a. de paclobutrazol	102,316	2,819	1,238
3	'Prata Anã' + 1,5g de i.a. de paclobutrazol	94,076	2,948	1,32
3	'Prata Anã' + 2,0g de i.a. de paclobutrazol	95,961	2,982	1,334
3	'FHIA 01' + 2,0g de i.a. de paclobutrazol	103,95	2,974	1,317
Média(G3)		99,076	2,931	1,302



REFERÊNCIAS

ALVAREZ, I.A.; CASTRO, P.R.C. Crescimento da parte aérea de cana crua e queimada. *Scientia Agricola*, v.56, p.1069-1079, 1999. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-901619990005000008&lng=en&nrm=iso>. Doi: 10.1590/S0103-901619990005000008.

BARROS, N.F. et al. Análise do crescimento de *Eucalyptus saligna* em solo de cerrado sob diferentes níveis de N.P.K. no Vale do Jequitinhonha, MG. *IPEF*, v.26, p.13-17, 1984.

BRODY, S. *Bioenergetics and growth*. New York: Reinhold, 1945. 645p.

ENGEL, V.L.; POGGIANI, F. Influência do sombreamento sobre o crescimento de mudas de algumas essências nativas e suas implicações ecológicas e silviculturais. *IPEF*, v.43/44, p.1-10, 1990.

HERNÁNDEZ, M.S. et al. Behavior of arazá (*Eugenia stipitata* Mc Vaugh) fruit quality traits during growth, development and ripening. *Scientia Horticulturae*, v.111, p.220-227, 2007. Disponível em: <http://www.sciencedirect.com/science?_ob=MIimg&_imagekey=B6TC3-4MGVJD3-3-F&_cdi=5159&_user=10&_orig=search&_coVerDate=02%2F05%2F2007&_sk=998889996&view=c&wchp=dGLbVtz-zSkWb&md5=ae3722c15110312407d3a57aedfce59b&ie=/sdarticle.pdf>. Doi:10.1016/j.scienta.2006.10.029.

KHATTREE, R.; NAIK, D.N. *Multivariate data reduction and discrimination with SAS software*. Cary, NC: SAS Institute, 2000. 588p.

NELDER, J.A. The fitting of a generalization of the logistic curve. *Biometrics*, v.17, p.89-94, 1961.

OLIVEIRA, H.N. et al. Comparação de modelos não-lineares para descrever o crescimento de fêmeas da raça Guzerá. *Pesquisa Agropecuária Brasileira*, v.35, p.1843-1851, 2000. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-204X2000000900017&lng=en&nrm=iso>. Doi: 10.1590/S0100-204X2000000900017.

REGAZZI, A.J. Teste para verificar a igualdade de parâmetros e a identidade de modelos de regressão não-linear. *Ceres*, v.50, p.9-26, 2003.

REGAZZI, A.J.; SILVA, C.H.O. Teste para verificar a igualdade de parâmetros e a identidade de modelos de regressão não-linear. I. Dados no delineamento inteiramente casualizado. *Revista de Matemática e Estatística*, v.22, p.33-45, 2004. Disponível em: <http://www.fcav.unesp.br/RME/fasciculos/v22/v22_n3/A3_Adair.pdf>. Acesso em: 04 Abr. 2007.

REZENDE, D.M.L.C. et al. Ajuste de modelos de platô de resposta para a exigência de zinco em frangos de corte. *Ciência e Agrotecnologia*, v.31, p.468-478, 2007. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-70542007000200030&lng=en&nrm=iso>. Doi: 10.1590/S1413-70542007000200030.

SANTORO, K.R. et al. Herdabilidade de parâmetros de curvas de crescimento não-lineares em zebuínos, no estado de Pernambuco. *Revista Brasileira de Zootecnia*, v.34, p.2280-2289, 2005. Disponível em: <<http://www.scielo.br/>

- scielo.php?script=sci_arttext&pid=S1516-35982005000700014&lng=pt&nrm=iso>. Doi: 10.1590/S1516-35982005000700014.
- SAS INSTITUTE. **Sas user's guide: statistics**. Cary, 2003. 846p.
- SCHWARZ, G. "Estimating the dimension of a model". **Annals of Statistics**, v.6, p.461-464, 1978.
- SHARMA, S. **Applied multivariate techniques**. New York: John Wiley & Sons, 1996. 493p.
- SOUZA, E.F.M.S. **Desenvolvimento das bananeiras 'Prata Anã' e 'FHIA 01' sob efeito do paclobutrazol aplicado no solo e nas folhas**. 2007. 61f. Dissertação (Mestrado em Fitotecnia) – Curso de Pós-graduação em Fitotecnia, Universidade Federal de Viçosa.
- URCHEI, M.A. et al. Análise de crescimento de duas cultivares de feijoeiro sob irrigação, em plantio direto e preparo convencional. **Pesquisa Agropecuária Brasileira**, v.35, n.3, p.497-506, 2000. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-204X2000000300004&lng=en&nrm=iso>. Doi: 10.1590/S0100-204X2000000300004.
- von BERTALANFFY, L. Quantitative laws in metabolism and growth. **Quarterly Review of Biology**, v.32, p.218, 1957.
- WINSOR, C.P. The Gompertz curve as a growth curve. **Proceedings of the National Academy of Science**, v.18, p.1-17, 1932.