

Ajuste do modelo logístico na descrição do crescimento de frutos de coco anão por meio de algoritmos iterativos

Thalita Kelen Leal do Prado¹

Taciana Villela Savian²

Joel Augusto Muniz¹

Adriana Matheus da Costa Sorato¹

1 Introdução

O coqueiro da variedade anão apresenta desenvolvimento vegetativo lento, é precoce, iniciando a produção em média com dois a três anos após o plantio. Chega a atingir 10 a 12m de altura e tem vida útil em torno de 30 a 40 anos. Apresenta estipe delgada, folhas numerosas, porém curtas e produz um grande número de pequenos frutos (150 a 200 frutos/planta/ano). (ARAGÃO, 2007).

Entre os modelos estatísticos utilizados para descrever o crescimento destaca-se por suas propriedades de ajuste o logístico (MAZZINI et al. 2005). De forma semelhante aos modelos lineares, o processo de estimação dos parâmetros em modelos não lineares pode ser feito utilizando-se de diferentes métodos, a saber: métodos de mínimos quadrados, método da máxima verossimilhança e método de Monte Carlo via cadeias de Markov-MCMC, considerando-se diferentes estruturas de erros.

Para a utilização do método MCMC é preciso obter as distribuições marginais que geralmente são complexas, necessitando de algoritmos iterativos como o de Gibbs Sampler (G-S) (CASELLA e GEORGE, 1992) e Metropolis-Hastings (M-H) (CHIB e GREENBERG, 1995), que são casos especiais do método de Monte Carlo via cadeias de Markov. No processo de cálculo envolvido nos algoritmos é necessário avaliar a convergência das cadeias. Alguns métodos, ou estratégias, são utilizados para monitorar a convergência das cadeias geradas por estes algoritmos sendo eles classificados como métodos informais e formais. Como método de monitoração formal tem-se, dentre os mais utilizados, aqueles propostos por GELMAN e RUBIN (1992), GEWEKE (1992), HEIDELBERGER e WELCH (1993) e RAFTERY e LEWIS (1992).

¹ DEX – UFPA. e-mail: tkprado1@yahoo.com.br

² DCE – ESALQ/USP.

^{1,2} Agracimento a Fapemig

O objetivo do trabalho foi avaliar o ajuste do modelo logístico, a dados do diâmetro externo longitudinal de frutos de coqueiros anão verde considerando a estrutura de erros autorregressiva, com base na amostragem da função de verossimilhança realizada pelo algoritmo MCMC e validar a metodologia utilizada no procedimento de ajuste por meio de um estudo de simulação baseado nos valores reais.

2 Material e métodos

Os dados utilizados para o ajuste do modelo foi obtido de um experimento (BENASSI, 2006) em que foi mensurado, dentre outras características, o diâmetro externo longitudinal (DEL), de frutos de coqueiro da variedade anão verde. A primeira medição foi feita no dia da abertura da inflorescência, e as demais, a cada 15 dias nos frutos marcados, totalizando 26 medições.

Foram obtidas 26 observações simuladas do diâmetro externo longitudinal (DEL) e os valores paramétricos utilizados para o estudo de simulação foram assumidos com base nos valores dos parâmetros obtidos por meio do método de mínimos quadrados.

O modelo logístico utilizado para estimar os parâmetros da curva de crescimento é dado por:

$$W_i = \frac{A}{1+B[\exp(-kt_i)]} + u_i \quad (1)$$

em que: W_i expressa o diâmetro (cm) do fruto na i -ésima medição; A é o parâmetro assintótico do modelo que corresponde ao valor máximo do diâmetro em relação ao tempo; B e k são parâmetros proporcionais à taxa de expansão do crescimento do fruto, t_i refere-se ao tempo na i -ésima medição, dado em dias após a abertura da inflorescência; $u_i = \phi_1 u_{i-1} + \dots + \phi_p u_{i-p} + \varepsilon_i$, com $i=1, \dots, n$ e n o número de medições no fruto; u_i é o resíduo do ajuste na i -ésima medição; ϕ_1 é o parâmetro autorregressivo de ordem 1; u_{i-1} é o resíduo do ajuste no tempo imediatamente anterior a i -ésima medição; ϕ_p é o parâmetro autorregressivo de ordem p ; u_{i-p} é o resíduo do ajuste em p tempos anterior a i -ésima medição; ε_i é o ruído branco com distribuição normal $N(0, \sigma^2)$. Para a estrutura de erros independentes tem-se $u_i = \varepsilon_i$. A existência de autocorrelação residual foi avaliada por meio do teste de Durbin-Watson (DW) e os resultados dos testes para autocorrelação residual orientaram a maneira de se construir a função de verossimilhança, e que foi posteriormente amostrada por meio do MCMC utilizando o software R (R Development Core Team, 2012).

3 Resultados e discussões

Para o ajuste do modelo aos dados reais do diâmetro longitudinal externo longitudinal considerando a estrutura de erros independentes (DEL sem AR), o valor da estatística DW foi de aproximadamente 0,64 para o modelo logístico. O valor é significativo ($p < 0,0001$) para o teste de autocorrelação positiva, indicando a necessidade de considerar um processo autorregressivo (AR), em princípio de primeira ordem, para a série de resíduos. Não houve necessidade de se considerar um processo autorregressivo de segunda ordem (AR2), de acordo com o teste, e foi adotada a precisão $\tau = 1/\sigma^2$ para efetuar os cálculos das marginais.

Para o estudo de dados simulados (Tabela 1) e reais (Tabela 2), após simular uma cadeia de tamanho 10.000, o critério de RAFTERY e LEWIS (1992), sugeriu 18.000 e 50.000 iterações para o tamanho da cadeia, *bunr-in* de 20 e 90 e um passo (*thin*) de 5 e 10, respectivamente, para os dados simulados e reais.

Para o critério de GEWEKE (1992), que consiste em indicar a convergência da média, resultou em um valor p sempre maior que o nível de significância adotado de 5%, não apresentando evidências contra a convergência. Segundo o critério de HEIDELBERGER e WELCH (1993), a hipótese nula de estacionariedade da amostra gerada não foi rejeitada para todos os parâmetros do modelo logístico com AR(1) (Tabela 1 e 2).

Por meio da tabela 1, observa-se uma simetria da distribuição para os parâmetros A e k. Pode-se observar que tanto a média como a moda podem ser consideradas como os valores que maximizam a função de verossimilhança. Para os parâmetros B, τ e ϕ obtém-se distribuições assimétricas, considerando a moda como o melhor valor estimado.

Tabela 1 Valores paramétricos (VP), Média, Moda, Erro Monte Carlo (EMC), Intervalo de Máxima Densidade (HPD) e o Critério de Geweke (valor p) das cadeias amostradas, para os dados simulados.

Parâmetros	VP	Média	Moda	EMC	HPD 95%		Geweke (Valor p)
					LI	LS	
A	22,00	21,50	21,51	$1,38 \cdot 10^{-2}$	20,30	22,47	0,6805
B	12,00	13,90	14,18	$1,03 \cdot 10^{-1}$	10,07	17,05	0,5787
k	0,02	0,019	0,019	$4,90 \cdot 10^{-7}$	0,019	0,020	0,5242
τ	0,80	0,99	0,97	$1,55 \cdot 10^2$	0,10	1,66	0,1575

ϕ	0,50	0,43	0,40	$9,87.10^{-3}$	-0,004	0,97	0,4713
--------	------	------	------	----------------	--------	------	--------

LI: limite inferior e LS: limite superior.

Os resultados apresentados para o Erro Monte Carlo (EMC) foram de baixa magnitude (Tabela 1 e 2), garantindo a convergência das cadeias geradas para cada parâmetro. Na Tabela 2, pode-se observar uma simetria da distribuição para os parâmetros A, B e k, desta forma, tanto a média como a moda podem ser consideradas como os valores que maximizam a função de verossimilhança. Para os parâmetros τ e ϕ obteve-se uma distribuição assimétrica adotando-se a moda como o melhor valor estimado.

Tabela 2 Média, Moda, Erro Monte Carlo (EMC), Intervalo de Máxima Densidade (HPD) e o Critério de Geweke (valor p) das cadeias amostradas, para os dados reais.

Parâmetros	Média	Moda	EMC	HPD 95%		Geweke (valor p)
				LI	LS	
A	20,0398	20,0622	$1,72.10^{-2}$	19,3384	20,7137	0,9736
B	21,2794	21,2472	$1,19.10^{-1}$	17,6069	23,9234	0,9916
k	0,0199	0,0199	$5,37.10^{-7}$	0,0199	0,0200	0,6624
τ	1,3994	1,3178	$8,19.10^{-3}$	0,6657	2,2047	0,8258
ϕ	0,6519	0,8521	$3,98.10^{-3}$	0,1565	0,9999	0,9663

LI: limite inferior e LS: limite superior.

No presente estudo, o diâmetro externo longitudinal estabiliza-se com um valor aproximado de 20,04 cm, após a abertura da inflorescência. SILVA et al. (2009) estudando características biométricas de frutos de coqueiro anão verde, obteve um valor de 17,23 cm para o diâmetro longitudinal, no sistema de produção convencional.

4 Conclusões

A metodologia utilizada no procedimento de ajuste por meio do estudo de simulação dos dados de diâmetro externo longitudinal (DEL) mostrou-se eficiente, considerando a estrutura de erro autorregressiva dos resíduos, apresentando estimativas bem próximas ao valor verdadeiro. O ajuste do modelo logístico aos dados experimentais de diâmetro externo longitudinal (DEL), considerando a estrutura autorregressiva de primeira ordem para os

resíduos, é adequado na descrição dos dados e o diâmetro máximo do coco anão estabiliza no valor entre 19,30 e 20,70 cm, com credibilidade de 95%.

5 Referências

- [1] ARAGÃO, W. M. Cultivares de coqueiros. In: FONTES, H. R.; FERREIRA, J. M. S.; SIQUEIRA, L. A. (Ed.). **A cultura do coqueiro**. Aracaju: Embrapa Tabuleiros Costeiros, 2007. (Embrapa Tabuleiros Costeiros. Sistemas de Produção, 1).
- [2] BENASSI, A. C. **Caracterizações biométricas, química e sensorial de frutos de coqueiro variedade anã verde**. 2006. 49 f. Tese (Doutorado em Produção Vegetal) - Universidade Estadual Paulista “Júlio de Mesquita Filho”, Jaboticabal, 2006.
- [3] CASELLA, G.; GEORGE, E. I. Explaining the Gibbs Sampler. **The American Statistician**. v. 46, n. 3, p. 167-174, 1992.
- [4] CHIB, S.; GREENBERG E. Understanding the Metropolis-Hastings Algorithm. **The American Statistician**. v. 49, n. 4, p. 327-335, 1995.
- [5] GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences. **Statistical Science**. Hayward, v. 7, n. 4, p. 457-472, 1992.
- [6] GEWEKE, J. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In: BERNARDO, J. M. et al. (Ed.). **Bayesian statistics**. New York: Oxford University, 1992. p. 625-631.
- [7] HEIDELBERG, P.; WELCH, P. Simulation run length control in the presence of an initial transient. **Operations Research**. Baltimore, v. 31, n. 6, p. 1109-1114, 1983.
- [8] MAZZINI, A. R. de A. et al. Curva de crescimento de novilhos Hereford: heterocedasticidade e resíduos autorregressivos. **Ciência Rural**. Santa Maria, v. 35, n.2, p. 422-427, 2005.
- [9] RAFTERY, A. L.; LEWIS, S. Comment: one long run with diagnostics: implementation strategies for Markov chain Monte Carlo. **Statistical Science**. Hayward, v. 7, n. 4, p. 493-497, 1992.
- [10] SILVA, D. L. V. da et al. Características físicas, físico-químicas e sensoriais da água de frutos de coqueiro anão verde oriundo de produção convencional e orgânica. **Ciência e Agrotecnologia**. Lavras, v. 33, n. 4, p. 1079-1084, 2009.
- [11] R DEVELOPMENT CORE TEAM. **R**: a language and environment for statistical computing. 2012. Disponível em: <<http://www.R-project.org>>. Acesso em: 03 fev. 2014.