

Ajuste de curvas de crescimento para a produção de cana de açúcar

Wederson Leandro Ferreira^{1*}, Fabio Vieira da Silva Junior², Juracy Mendes Moreira³, Guido Gustavo Humada Gonzalez⁴, Juliano Bortolini⁵, Aurélio Ferreira Melo⁶

¹Instituto Federal Farroupilha – Campus Alegrete-RS E-mail: Wederson.ferreira@iffarropilha.edu.br *Autor para correspondência

²Instituto Federal Farroupilha – Campus Alegrete-RS E-mail: Fabio.silva@iffarropilha.edu.br

³Prof mestre. Instituto Federal Goiano. Campus Rio Verde - GO E-mail: juracimendesmoreira@yahoo.com.br

⁴Prof. Doutor. Universidad San Carlos - Asuncion- Paraguai E-mail: gustavohumad@hotmail.com

⁵Universidade Federal do mato Grosso - Departamento de Estatística UFMT E-mail: julianobortolini@ufmt.br

⁶Doutorando Instituto Federal Ciência e Tecnologia Goiano Campus Rio Verde – GO E-mail: aurelioferreiramelo@hotmail.com

RESUMO: A modelagem de curvas de crescimento é muito comum em dados provenientes de medidas repetidas no tempo, porque é possível estabelecer uma relação funcional entre o tempo e os tratamentos utilizados. O objetivo principal deste trabalho é obter uma relação funcional que estabeleça uma associação polinomial entre a produção, em toneladas, e os 5 anos em que 3 genótipos de cana de açúcar tiveram suas produções aferidas em um experimento instalado na Usina Porto Rico, em Alagoas, conduzido pela UFSC. O delineamento utilizado foi o blocos ao acaso, com 4 repetições. Nestes modelos assume-se curvas médias fixadas, em torno das quais existem variações aleatórias para as curvas individuais das respectivas unidades experimentais. Ficou constatado pelos critérios de AIC e BIC que a melhor estrutura é a componente da variância (VC) para representar a matriz de covariâncias do erro experimental e não estruturada (UN) para a matriz de covariância dos parâmetros de efeitos aleatórios e que o polinômio de 3º grau foi selecionado para representar a relação média entre produção e os 5 anos avaliados para os três genótipos.

Palavras chave: Análise de perfil, Cana de açúcar, dados Longitudinais

Adjustment of growth curves for the production of sugarcane

ABSTRACT. The modeling of growth curves is very common in data from repeated measures in time, because it is possible to establish a functional relationship between the time and the treatments used. The main objective of this work is to obtain a functional relationship that establishes a polynomial association between the production, in tons, and the 5 years in which 3 sugarcane genotypes had their production measured in an experiment installed in the Puerto Rico Plant, in Alagoas, conducted by UFSC. The experimental design was the randomized blocks, with 4 replicates. In these models we assume fixed mean curves, around which there are random variations for the individual curves of the respective experimental units. It was verified by the AIC and BIC criteria that the best structure is the variance component (VC) to represent the covariance matrix of the experimental and unstructured error (UN) for the covariance matrix of the random effects parameters and that the polynomial of 3rd degree was selected to represent the average relation between production and the 5 years evaluated for the three genotypes.

Keywords: Profile analysis, sugarcane, longitudinal data

Ajuste de curvas de crecimiento para la producción de caña de azúcar

RESUMEN. El modelado de curvas de crecimiento es muy común en datos provenientes de medidas repetidas en el tiempo, porque es posible establecer una relación funcional entre

el tiempo y los tratamientos utilizados. El objetivo principal de este trabajo es obtener una relación funcional que establezca una asociación polinomial entre la producción en toneladas y los 5 años en que 3 genotipos de caña de azúcar tuvieron sus producciones aferidas en un experimento instalado en la Usina Puerto Rico, Alagoas, conducido por la UFSC. El delineamiento utilizado fue el bloque al azar, con 4 repeticiones. En estos modelos se asumen curvas medias fijadas, alrededor de las cuales existen variaciones aleatorias para las curvas individuales de las respectivas unidades experimentales. Se constató por los criterios de AIC y BIC que la mejor estructura es el componente de la varianza (VC) para representar la matriz de covariancias del error experimental y no estructurado (UN) para la matriz de covarianza de los parámetros de efectos aleatorios y que el polinomio de los alelos 3º grado fue seleccionado para representar la relación media entre producción y los 5 años evaluados para los tres genotipos.

Palabras clave: Análisis de perfil, caña de azúcar y datos longitudinales

Introdução

No Brasil são desenvolvidos muitos experimentos com cana de açúcar devido ao crescente mercado (Freitas, 2008); sobretudo com o advento tecnológico proveniente dos carros flex. Grande número de genótipos é avaliado todos os anos, gerando dados de produtividade com os quais procuram se indicar novas variedades para cada região produtora.

Técnicas de análises de experimentos que utilizam medidas repetidas ao longo do tempo devem considerar a estrutura de correlação inter e intra-tempos. Neste trabalho foi considerado o fato relevante que, às vezes, não é estudado, ou seja, dados mensurados longitudinalmente podem ocasionar correlações as observações e consequentemente dependência dos erros. Os modelos de coeficientes aleatórios são casos particulares de modelos lineares mistos e possuem as mais variadas aplicações e tornaram-se mais conhecidos nas últimas décadas devido ao advento de técnicas computacionais robustas. Dentre essas aplicações destaca-se a modelagem de curvas de crescimento, principalmente de dados provenientes de medidas repetidas no tempo.

O objetivo principal deste trabalho foi obter a relação polinomial entre a produção, tonelada por hectare (TH), média, e o período de 5 anos, em que os três genótipos escolhidos no experimento tiveram suas produções aferidas.

Material e Métodos

Na Usina Porto Rico, em Alagoas, foi instalado um experimento conduzido pelos Programas de Melhoramento Genético da Cana de açúcar (PMGCA) da Universidade Federal de São Carlos e Universidade Federal de Alagoas (Freitas, 2008). O experimento foi instalado em 15/09/2001 na Fazenda Santa Tereza, unidade da Usina Porto

Rico, para avaliar vários genótipos, utilizando-se delineamento em blocos ao acaso, com quatro repetições. Cada parcela tinha 50,4m² de área e espaçamento de 1,40 m entre linhas. Conduzido no período de setembro de 2001 a outubro de 2006 e foram obtidas informações de cinco colheitas, analisado a variável TCH (Tonelada de Cana por Hectare) e através dessa informação foi estudado particularmente o perfil médio de respostas de três genótipos de cana de açúcar, escolhido entre os diferentes genótipos contidos no experimento.

No ajuste de curvas de crescimento, quando há interesse em encontrar uma relação entre a variável resposta e o tempo, os coeficientes de regressão das curvas individuais serão considerados efeitos aleatórios, variando de unidade experimental para unidade experimental, e os coeficientes de regressão das curvas médias, efeitos fixos (Hauser et al., 2009, Ferreira and Moraes, 2013).

A especificação do modelo linear misto pode ser feita na seguinte forma matricial (Henderson, 1975):

$$y = X\beta + Zu + e \quad (1)$$

em que, y representa o vetor das respostas individuais ou vetor de observações, de dimensões $n \times 1$; X é uma matriz do planejamento experimental que especifica os parâmetros de efeitos fixos, de dimensões $n \times r$, e posto $r < n$; β é o vetor de parâmetros fixos, desconhecidos, de dimensão $r \times 1$; Z é uma matriz conhecida do planejamento experimental, que especifica os efeitos aleatórios, de dimensões $n \times q$; u é o vetor de efeitos aleatórios, desconhecidos, $q \times 1$ e e é o vetor de erros aleatórios não observáveis, de dimensão $n \times 1$. As matrizes X e Z podem se diferenciar, sendo que Z pode conter qualquer covariável que influencie a unidade experimental.

A formulação de X é semelhante a da regressão. No modelo (1) assume-se que $E[u] = 0$ e $E[e] = 0$ com matrizes de covariâncias representadas por

$$V \begin{bmatrix} u \\ e \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix},$$

com suposição que u e e são *iid*, então $u \sim N(0, G)$ e $e \sim N(0, R)$, (Ferreira, 2011). O modelo de coeficientes aleatórios é um caso especial de modelo linear misto e é usualmente especificado em termos das respostas condicionais aos efeitos aleatórios, sendo y um vetor específico de medidas repetidas para alguma unidade experimental:

$$y|u \sim NM(X\beta + Zu, R).$$

Entretanto, as inferências são realizadas no modelo marginal, e pode-se demonstrar com as premissas acima que $y \sim NM(X\beta, ZGZ' + R)$ ou $y \sim NM(X\beta, V)$. Henderson (1963) provaram que $\hat{\beta}$, dado pelo modelo (1) coincide com a solução obtida pelo método de mínimos quadrados generalizados, cuja expressão é $\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}y$ e a predição de \hat{u} , segundo Henderson (1963) é dado por $\hat{u} = GZV^{-1}(y - X\hat{\beta})$.

Resultados e Discussão

A escolha dos efeitos aleatórios foi realizada através da análise de regressão polinomial clássica para cada unidade experimental, no qual ficou constatada que o intercepto, componente linear, quadrático e o cúbico podem ser considerados como aleatórios, pois houve confirmação da significância destes parâmetros em todas as unidades experimentais avaliadas. Já a escolha da curva polinomial média ajustada para a produção média foi realizada através da análise de regressão polinomial, entretanto, neste caso, considerando o peso médio para cada genótipo e sendo o polinômio de 3º grau escolhido, devido ao bom ajuste indicado pela análise de regressão.

Atribuindo as estruturas de covariâncias simetria composta (CS), componente de variância (VC), não estruturada (UN) e a de Huynh-Feldt (H-F) para a escolha da estrutura mais adequada para os efeitos aleatórios e também do erro experimental do modelo (1), realizado pelo proc mixed do SAS, ficou constatado pelos critérios de AIC e BIC que a melhor estrutura é não

estruturada (VC) para R e auto não estruturada (UN) para G , em função de possuir menores valores (Tabela 1).

Tabela 1. Estimativas dos critérios de seleção das estruturas da matriz de covariâncias para G , variação entre indivíduos, e R , variação intra-indivíduo, utilizando o método de estimação da máxima verossimilhança restrita (MVR)

Estruturas		Estruturas	
R	G	AIC	BIC
VC	UN	436,0	438,5
VC	CS	438,5	439,5
VC	H-F	438,5	439,8
UM	UN	468,0	478,5
UM	CS	448,5	458,5
UM	H-F	458,5	468,5

O teste F, tabela 2, aplicado aos efeitos fixos do modelo indicaram que todos os efeitos são significativos a 5% de probabilidade, com exceção da fonte de variação bloco.

Tabela 2. Testes para efeitos fixos do modelo utilizando as estruturas de covariância UN para G e VC para R para o ajuste do modelo da produção dos 3 genótipos de cana de açúcar.

Causas de variação	de GL/num.	GL/de	F	Pr>F
Bloco	3	27	0,44	0,72
Genótipos	2	27	11,24	<0,01
Anos	1	9	16,25	<0,00
Gen*Anos	2	9	7,71	<0,02
Gen * Anos^2	3	9	5,78	<0,02
Gen * Anos^3	3	9	5,94	<0,02

Após ter selecionado a estrutura UN, efeitos aleatórios, VC para o erro do modelo, procede-se a última etapa que é a estimação dos parâmetros das curvas polinomiais de crescimento do peso médio da produção.

Na Figura 1, encontram-se a representação gráfica e as estimativas das equações polinomiais de terceiro grau para os três genótipos. Observa-se, em média, que o genótipo RB92579 obteve maior ganho de produção, comparado aos genótipos SP791011.

As curvas polinomiais ajustadas para os genótipos RB92579, SP791011, RB835994 e RB92579 são, respectivamente, $\hat{y} = 152,4 - 78,8t + 32,8t^2 - 4t^3$, $\hat{y} = 114,4 - 141,8t + 51t^2 - 5,7t^3$ e $\hat{y} = 116,6 - 171,5t + 61t^2 - 6,7t^3$.

As curvas de crescimento (Figura 1) ajustadas para os genótipos SP791011, RB835994 e RB92579 são, respectivamente:

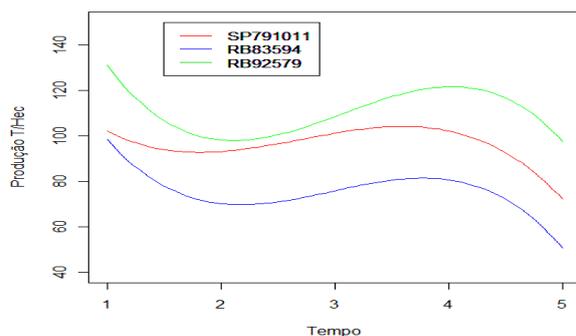


Figura 1. Curvas da produção de cana de açúcar (2001-2006) dos três genótipos estudados.

Por fim, o diagnóstico da qualidade do ajuste da modelagem das curvas polinomiais para a produção dos genótipos pode ser visualizado na Figura 2, a qual mostra uma tendência linear dos resíduos, não notando valores atípicos e também pela aplicação do teste de Shapiro-Wilks em que foi obtido $w = 0,9797$, com valor $p = 0,5662$, confirmado a normalidade dos resíduos.

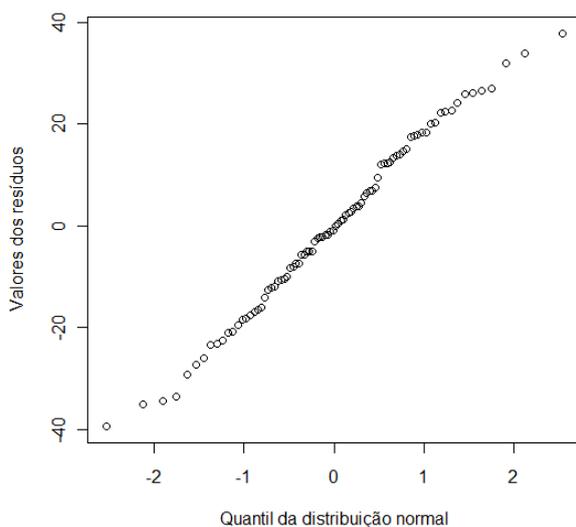


Figura 2. Representação gráfica da probabilidade normal dos resíduos das curvas ajustadas à produção, THC, dos genótipos de cana de açúcar.

Conclusões

Os resultados dos critérios informação de Akaike (AIC) e o de Schwars (BIC), no que tange a seleção do modelo mais adequado indicaram a

seleção das mestras estruturas de covariâncias. As curvas médias estimadas da produção apresentaram bom ajuste ao explicar a produção média dos genótipos de cana de açúcar no desencadear dos anos avaliados, optando pela curva de 3º grau como a mais adequada.

Referências Bibliográficas

- Ferreira, D. F. 2011. SISVAR: A Computer Statistical Analysis System. *Ciência e Agrotecnologia*, 35, 1039-1042.
- Ferreira, W. L. & Moraes, A. R. 2013. Análise da influência do café no ganho de peso de animais (ratos) por meio de modelo linear misto. *Revista Brasileira de Biometria*, 31, 485-500.
- Freitas, E. G. 2008. Análise de dados longitudinais em experimentos com cana-de-açúcar. *Departamento de Agronomia*. Universidade de São Paulo, Piracicaba, São Paulo.
- Hauser, L., Vigo, Á., Kieling, C., Jardim, L. B., Camey, S. A. & Leotti, V. B. 2009. Curva de crescimento usando modelo misto: Uma aplicação na progressão da doença de Machado-Joseph. *Clinical & Biomedical Research*, 29, 5-17.
- Henderson, C. R. 1963. Selection index and expected genetic advance. *Statistical genetics and plant breeding*, 982, 141-163.
- Henderson, C. R. 1975. Best linear unbiased estimation and prediction under a selection model. *Biometrics*, 31, 423-447.

Article History:

Received 24 April 2017

Accepted 16 May 2017

Available on line 17 July 2017

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License 4.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.