

Seleção da matriz de variância-covariância residual na análise de ensaios varietais com medidas repetidas em cana-de-açúcar

Selection of the residual variance-covariance matrix in the analysis of varietal trials with repeated measures in sugarcane

Emerson Noletto Silva^I João Batista Duarte^{II} Américo José dos Santos Reis^{III}

RESUMO

Este trabalho objetivou avaliar diferentes estruturas da matriz de variâncias e covariâncias residual (Σ), quanto ao ajustamento de dados longitudinais via modelos mistos, em experimentos varietais de cana-de-açúcar. A seleção adequada desta matriz garante a escolha de um modelo mais representativo dos dados. Em cada modelagem, variou-se ainda a suposição associada aos efeitos de tratamentos (variedades), como fixos e aleatórios. Quatro ensaios varietais, conduzidos entre 2005 e 2009, em três localidades do Estado de Goiás, foram considerados. Cada experimento foi delineado em blocos casualizados com três ou quatro repetições. A variável resposta analisada foi toneladas de colmos por hectare (TCH). Para avaliar a qualidade de ajustamento dos modelos, foram utilizados o critério de informação de Akaike (AIC) e o teste da razão de verossimilhanças. Este último foi utilizado apenas para comparar modelos hierárquicos, tomados dois a dois. Constatou-se que as análises pelo modelo univariado clássico de parcelas subdivididas oscilaram entre as piores ou entre aquelas de ajustes apenas medianos. As estruturas da matriz Σ com os melhores ajustamentos variaram entre os ensaios, com destaque para a matriz não-estruturada. Tais resultados revelam que a estrutura de erros independentes, em geral, não se mostra adequada para esse tipo de análise e, também, que não é seguro definir previamente uma estrutura específica de Σ para tais ensaios. Pequenas alterações foram observadas na classificação das estruturas ao se assumirem os efeitos de tratamentos como fixos ou aleatórios; porém, sem efeito importante na classificação das melhores estruturas em cada ensaio.

Palavras-chave: *Saccharum* spp., dados longitudinais, estruturas de covariância, modelos mistos, genótipos aleatórios.

ABSTRACT

This study aimed to evaluate different residual structures of variance-covariance matrix (Σ), regarding the fitting

of longitudinal data via mixed models in variety trials of sugarcane. The adequate choice of this matrix provides most representative models to the data. In each model was also evaluated the effects of treatments (varieties), either as fixed or as random. Four trials were carried out in three locations in the Goiás State, Brazil, from 2005 to 2009. Each experiment was designed in randomized complete block with three or four repetitions. The response variable analyzed was tons of stalks per hectare (TCH). The goodness of fitting of the different models to the data was assessed by Akaike information criterion (AIC) and by likelihood ratio test (LRT). This last statistic was used only to compare nested models, two by two. It was observed that classic model in split-plot design ranged among the worst or with just median adjustments. The structures of Σ matrix with the best fittings to the data varied among trials, with outstanding for the unstructured matrix. These results show that the structure of independent errors, in general, is not adequate for these analyses, and a prior definition of the co-variance structure can lead to unreliable results for these trials. Small changes were observed in the ranking of these structures by assuming the treatment effects as fixed or random, however, without significant effects on the ranking of the best structures in each trial.

Key words: *Saccharum* spp., longitudinal data, covariance structures, mixed models, random genotypes.

INTRODUÇÃO

A avaliação de clones experimentais em cana-de-açúcar comumente envolve a amostragem na mesma planta ou parcela ao longo dos cortes (cana-planta, cana-soca, ressoça etc). Esse fato é de grande importância para avaliar as relações entre os diferentes cortes durante o desenvolvimento

^IPrograma de Pós-graduação em Agronomia (PPGA), Universidade Federal de Goiás (UFG), Goiânia, GO, Brasil.

^{II}Setor de Melhoramento de Plantas, Escola de Agronomia, UFG, CP 131, 74001-970, Goiânia, GO, Brasil. E-mail: jbduarte@ufg.br. Autor para correspondência.

^{III}Rede Interuniversitária para o Desenvolvimento do Setor Sucroenergético (Ridesa Brasil), Programa de Melhoramento Genético da Cana-de-Açúcar (PMGCA), UFG, Goiânia, GO, Brasil.

dos clones. Nesse tipo de experimento, ocorre interdependência entre as observações, devido às medidas repetidas nos diferentes tempos sobre a mesma parcela, o que dificulta a análise estatística dos dados. Para contornar este problema, podem-se analisar as observações (dados longitudinais) através de diferentes estruturas da matriz de variâncias-covariâncias residuais (Σ). Porém, a apresentação adequada das observações depende de seu ajustamento ao modelo de estrutura selecionado (FREITAS et al., 2008).

Na busca do modelo estatístico que melhor represente os dados, levando em consideração se estes são independentes ou não, várias estruturas de matriz de variâncias-covariâncias (Σ) precisam ser avaliadas (QUINTAL, 2013). Para esta finalidade, uma alternativa é utilizar a abordagem de modelos mistos, que possibilita contemplar a estrutura de interdependência das observações, acomodando aquela mais apropriada à correlação presente nas medidas repetidas, tomadas na mesma unidade experimental.

Várias técnicas de seleção do modelo com a respectiva estrutura da matriz de covariâncias podem ser utilizadas. As de uso mais difundido são os critérios de informação de Akaike – AIC (Akaike's Information Criterion) e o de Schwarz – BIC (Bayesian Information Criterion); ambos baseados na verossimilhança de ajuste do modelo e dependentes do número de observações e parâmetros (FLORIANO et al., 2006). Em geral, os dois critérios produzem resultados concordantes. Outro procedimento de uso comum é o teste assintótico de razão de verossimilhanças (LRT), que permite comparar dois modelos de cada vez, ambos ajustados por verossimilhança, um deles como versão restrita do outro (modelos aninhados ou encaixados). O uso dessas técnicas é fundamental na teoria de decisão em modelo misto, pois, além da qualidade de ajustamento, consideram o princípio da parcimônia, que penaliza modelos com maior número de parâmetros (CAMARINHA FILHO, 2002).

Nesse contexto, o objetivo deste estudo foi comparar modelos estatísticos com diferentes estruturas de matriz de variâncias-covariâncias (Σ) para a análise de dados com medidas repetidas em experimentos varietais de cana-de-açúcar. Buscou-se também avaliar a influência da suposição associada aos efeitos de tratamentos genéticos (variedades), como fixos ou aleatórios, sobre os resultados dos ajustes obtidos sob diferentes estruturas de Σ .

MATERIAL E MÉTODOS

O material de pesquisa consistiu num grupo de quatro experimentos varietais de cana-de-açúcar, conduzidos entre 2004 e 2009, em três localidades do Estado de Goiás: Goiatuba (usina Goiasa, com dois ensaios), Goianésia (usina Jalles Machado) e Jandaia (usina Denusa). Os ensaios foram instalados no delineamento de blocos completos casualizados, com quatro (Goiatuba e Jandaia) ou três repetições (Goianésia). Os clones ou tratamentos foram provenientes do Programa de Melhoramento Genético da Cana-de-açúcar da Universidade Federal de Goiás (PMGCA-UFG), vinculado à Rede Interuniversitária para o Desenvolvimento do Setor Sucroenergético (Ridesa Brasil). Originaram-se de cruzamentos realizados em 1995 e 1996, razão por que são referidos como clones das Séries 95 e 96.

As parcelas foram constituídas de cinco fileiras de plantas, de 10 m, com espaçamento de 1,5m entre fileiras. No plantio, o número de gemas por metro de sulco variou entre oito a vinte. Neste estudo, foram considerados apenas os dados da variável tonelada de cana-de-açúcar por hectare (TCH), nos diferentes cortes (quatro nos ensaios de Goiatuba e Goianésia, e três no ensaio de Jandaia). Em cada experimento, as avaliações foram realizadas sobre a mesma parcela, caracterizando o delineamento como “parcelas subdivididas no tempo” ou “ensaio de medidas repetidas”.

A análise dos dados via abordagem de modelos mistos teve como objetivo avaliar estruturas da matriz de variâncias-covariâncias residuais diferentes de $\Sigma=I\sigma^2$ (modelo clássico, de erros independentes), quanto à sua adequação aos dados. Assim, conforme descreve FREITAS (2007), procedeu-se à análise mediante ajustamento do seguinte modelo:

$$y_i = X_i \beta + Z_i \gamma_i + e_i$$

em que: y_i é o vetor coluna ($n_i \times 1$) das n_i observações tomadas da unidade experimental “i” ao longo do tempo ou condição de avaliação; β é o vetor ($p \times 1$) de parâmetros fixos desconhecidos, em que a dimensão ‘p’ é o número de níveis de efeitos fixos; X_i é a matriz ($n_i \times p$) de incidências associadas aos elementos do vetor β ; γ_i é o vetor ($g \times 1$) de efeitos aleatórios desconhecidos, assumindo-se $\gamma_i \sim N(0, G_i = I\sigma_g^2)$, em que σ_g^2 é variância genotípica comum (este efeito apenas esteve presente no modelo quando se assumiu genótipos de efeitos aleatórios); Z_i é a matriz ($n_i \times g$) de incidências associadas aos elementos do vetor γ_i (idem observação para γ_i , com respeito à suposição

de aleatoriedade para tais efeitos); e_i é o vetor ($n_i \times 1$) de erros aleatórios, com $e_i \sim N(0, R_i = \Sigma)$, em que Σ é a matriz de variâncias-covariâncias residuais, que, neste caso, assumiu diferentes estruturas.

As estruturas de matriz de variâncias-covariâncias residuais ($R=\Sigma$) avaliadas, em conformidade com aplicações usuais na literatura de análise de medidas repetidas (FREITAS, 2007; QUINTAL, 2013) foram: parcela subdividida clássica (PSC) – estrutura de erros independentes baseada no ajuste do modelo univariado clássico; componente de variância (VC) – variâncias desiguais entre cortes ($\sigma^2_i; i=1,2,\dots,c$ cortes) e observações independentes; simetria composta (CS) – variâncias ($\sigma^2 + \sigma_1$) e covariâncias (σ_1) homogêneas; autoregressiva de primeira ordem (AR(1)) – variâncias homogêneas (σ^2) e covariâncias reduzindo-se exponencialmente à medida que se aumenta o intervalo de tempo ($\sigma^2\rho, \sigma^2\rho^2, \sigma^2\rho^3,\dots$; em que ρ , a correlação entre cortes, é o parâmetro autorregressivo assumido $|\rho|<1$ em processo estacionário); *Toeplitz* (TOEP) – variâncias homogêneas (σ^2) e covariâncias desiguais (s_{ij} ; cortes $i \neq j$); simetria composta heterogênea (CSH) – variâncias desiguais (σ^2_i) e covariâncias entre cortes dadas por $\sigma_i\sigma_j\rho$; *Huynh-Feldt* (HF) – variâncias desiguais (σ^2_i) e covariâncias dadas por $\frac{1}{2}(\sigma^2_i + \sigma^2_j) - \lambda$ (em que λ é a diferença entre a média das variâncias e a média das covariâncias); autoregressiva de primeira ordem heterogênea (ARH(1)) – com variâncias desiguais (σ^2_i) e covariâncias reduzindo-se exponencialmente, conforme: $\sigma_i\sigma_j\rho, \sigma_i\sigma_j\rho^2,\dots$); não estruturada (UN) – todas as variâncias e covariâncias desiguais. Maior detalhamento sobre estas e outras estruturas da matriz R são disponíveis em SAS INSTITUTE (2012).

Para avaliar a qualidade do ajustamento proporcionado pelos modelos sob diferentes estruturas da matriz Σ , em conformidade com RESENDE et al. (2006), foram utilizados o critério de informação de Akaike (AIC) e o teste da razão de verossimilhanças (LRT). O critério AIC é baseado na teoria de decisão e, para evitar excesso de parametrização, penaliza os modelos com número grande de parâmetros. Assim, é definido pela expressão: $AIC = -2l + 2p$, em que l é o logaritmo neperiano da função de verossimilhança do modelo, e “ p ” é o número de parâmetros da matriz de variâncias-covariâncias (Σ). Segundo este critério, o modelo de matriz de variância-covariância a ser escolhido é aquele com o menor valor de AIC. Para compatibilizar a comparação dos valores de AIC de todos os modelos (com diferentes estruturas de Σ), o modelo univariado clássico foi ajustado sem o

termo associado ao “erro a” (em nível de parcelas). Embora sem levar em conta a parcimônia do modelo, a estatística $-2l$, também denotada “ -2 Reml ” (relativo ao método de estimação de variâncias por “máxima verossimilhança restrita”), representa uma medida adicional da qualidade de ajustamento; em que menor valor implica melhor ajuste.

O teste LRT é utilizado para comparar dois modelos estimados por máxima verossimilhança, sendo um deles uma versão restrita do outro (modelos aninhados ou encaixados); isto é, ‘modelo completo’ versus ‘modelo reduzido’, com o primeiro tendo ‘ r ’ parâmetros adicionais. A estatística do teste avalia se os parâmetros adicionais melhoram significativamente o modelo. Assim, considerando-se $L_1 = -2l$, para o modelo com menor número de parâmetros, e $L_2 = -2l$ para o modelo com maior número de parâmetros (com r parâmetros extras), a hipótese em teste é a de que os dois modelos são equivalentes; isto é, os parâmetros extras não diferem de zero. Sob normalidade, a diferença entre L_1 e L_2 é assintoticamente distribuída como: $L_1 - L_2 \sim \chi^2_{[r]}$ (qui-quadrado com r graus de liberdade).

As análises foram realizadas no programa estatístico SAS (SAS INSTITUTE, 2012), via procedimento para ajuste de modelos lineares mistos (Proc Mixed).

RESULTADOS E DISCUSSÃO

Antes de comparar a qualidade de ajustamento associado às diferentes modelagens, é importante destacar certos resultados preliminares. Um deles refere-se à correlação entre cortes nos diferentes ensaios, que se manifestou com tendência positiva, embora de baixa magnitude ($IC^{95\%}: 0,25 \pm 0,09$). O segundo corte foi o que se mostrou mais correlacionado com os outros ($r=0,32$). Assim, não se confirmou a hipótese comumente aventada de falta de correlação entre as produções em cana-planta e nos demais cortes ($r=0,28$), o que ratifica a dependência natural entre cortes. Demonstra-se, portanto, que a suposição de independência não pode ser admitida como regra para sustentar um modelo de análise de variância clássica para esses ensaios (COSTA, 2003; FREITAS et al., 2008; QUINTAL, 2013). Tal constatação ainda foi corroborada pelos resultados significativos ($P<0,01$) do teste de esfericidade (MAUCHLY, 1940), aplicado a cada ensaio. Tudo isso justificou a investigação por alternativas de análise que acomodassem correlações dessa natureza.

Observando-se os resultados do ajuste dos modelos em cada ensaio, naquele conduzido

em Jandaia, o modelo com matriz Σ não estruturada (UN) foi o que apresentou melhor ajustamento aos dados (menores valores de -2 Reml e AIC); tanto assumindo os efeitos genotípicos como fixos ou aleatórios (Tabela 1). Esta superioridade foi confirmada pela significância ($P < 0,05$) dos testes de razão de verossimilhanças, que compararam este modelo com os que imediatamente o sucederam nos valores de AIC (Tabela 2). A comparação estatística entre os demais pares sucessivos de modelos também ratifica a constatação de que a mudança na suposição dos efeitos genotípicos não afetou significativamente a sua ordem de classificação. Observa-se ainda que as estruturas PSC, VC e CS foram as mais inadequadas ao ajustamento dos dados deste ensaio, as quais não se diferenciaram ($P > 0,05$), seja na suposição de genótipos fixos ou aleatórios. Logo, não se recomendaria o modelo univariado clássico (PSC) para a análise deste conjunto de dados.

No primeiro dos ensaios em Goiatuba (Série 95 de cruzamentos), a classificação dos modelos com diferentes estruturas de Σ mostrou certa variação sob mudança na suposição dos efeitos genotípicos (Tabela 3). Isso, contudo, não afetou estatisticamente a eleição dos melhores e piores modelos, em ambos os casos. Tanto sob genótipos fixos, quanto sob genótipos aleatórios, as estruturas com os menores valores de AIC (melhor ajustamento aos dados) e que não diferiram estatisticamente entre si ($P > 0,05$) foram CSH (simetria composta heterogênea) e UN (não estruturada). O bom ajustamento associado à estrutura CSH também foi

observado por CECON et al. (2008). Já as piores estruturas (maiores valores de AIC), que também não diferiram entre si, embora de todas as demais, foram PSC, VC e AR(1).

Quanto às duas melhores estruturas, CSH e UN, vale destacar que, embora não se diferiram ($P > 0,05$), a primeira delas é vantajosa, pois requer menor número de parâmetros para garantir a mesma qualidade de ajustamento. Entre aquelas de desempenho intermediário, houve considerável mudança na classificação. Por exemplo, enquanto, na condição de genótipos fixos, a estrutura HF ficou entre as três melhores; na de genótipos aleatórios, esta ficou entre as de desempenho intermediário (Tabela 3). O mesmo aconteceu com a estrutura TOEP, ao se inverter tais suposições. Isso demonstra que a especificação inadequada acerca da suposição do efeito genotípico também pode induzir à tomada de decisão errônea sobre a escolha da melhor estrutura da matriz de variâncias-covariâncias residuais. Por último, deve-se destacar o mau desempenho de ajustamento proporcionado pela estrutura PSC, que, tal como no ensaio anterior, ratifica a não recomendação do modelo univariado clássico. Isso corrobora as recomendações de que seja essencial modelar a estrutura de covariâncias na análise de dados longitudinais, para se garantir inferências válidas (QUINTAL, 2013).

No outro ensaio conduzido em Goiatuba (Série 96), as duas suposições para os efeitos genotípicos (fixos ou aleatórios) também praticamente não alteraram a identificação da melhor

Tabela 1 - Resultados do ajustamento de modelos com diferentes estruturas de matriz de variâncias-covariâncias residuais (Σ), segundo o número de parâmetros (NP), valor máximo da função de verossimilhança (-2 Reml) e do critério de informação de Akaike (AIC), em análises de medidas repetidas de ensaio de competição de variedades de cana-de-açúcar, conduzido em Jandaia-GO (2005-2008), sob avaliação de três cortes e efeitos genotípicos assumidos como fixos ou aleatórios.

Genótipos fixos				Genótipos aleatórios			
Modelos	NP	-2 Reml	AIC	Modelos	NP	-2 Reml	AIC
UN	6	2248,601	2260,601	UN	8	2924,906	2940,906
ARH (1)	4	2255,703	2263,703	ARH (1)	6	2933,739	2945,739
CSH	4	2257,004	2265,004	CSH	6	2935,471	2947,471
HF	4	2263,618	2271,618	HF	6	2941,500	2953,500
AR (1)	2	2283,700	2287,700	AR(1)	4	2958,734	2966,734
TOEP	3	2283,735	2289,735	TOEP	5	2958,489	2968,489
CS	2	2286,397	2290,397	CS	4	2962,379	2970,379
VC	1	2289,700	2291,709	VC	3	2965,200	2971,200
PSC ¹	1	2289,709	2291,709	PSC	3	2965,200	2971,200

¹ Para obtenção de valores de AIC comparáveis, o modelo PSC (parcelas subdivididas), correspondente à estrutura de erros independentes, foi ajustado sem o termo relativo ao erro "a" (genótipos x blocos).

Tabela 2 - Comparação de modelos com diferentes estruturas de matriz de variâncias-covariâncias residuais (Σ), tomados dois a dois, segundo o teste da razão de verossimilhanças (χ^2), baseado em análises de medidas repetidas de experimento de competição de variedades de cana-de-açúcar, conduzido em Jandaia-GO (2005-2008), sob avaliação de três cortes e assumindo os efeitos genotípicos como fixos ou aleatórios.

-----Genótipos fixos-----				-----Genótipos aleatórios-----			
Estruturas	DNP ¹	χ^2	p-valor	Estruturas	DNP	χ^2	p-valor
UN-ARH(1)	2	7,110	0,028581	UN-ARH(1)	2	8,790	0,012339
UN-CSH	2	8,340	0,015452	UN-CSH	2	10,47	0,005327
UN-HF	2	14,94	0,000570	UN-HF	2	16,58	0,000251
UN-AR(1)	4	35,07	<0,00001	UN-AR(1)	4	33,78	0,000001
UN-TOEP	3	35,06	<0,00001	UN-TOEP	3	33,61	<0,00001
UN-CS	4	37,44	<0,00001	UN-CS	4	37,45	<0,00001
UN-VC	5	41,11	<0,00001	UN-VC	5	40,29	<0,00001
UN- PSC ²	5	41,11	<0,00001	UN- PSC	5	40,29	<0,00001
ARH(1)-AR(1)	2	27,66	<0,00001	ARH(1)-AR(1)	2	25,00	0,000004
ARH(1)-TOEP	1	27,94	<0,00001	ARH(1)-TOEP	1	24,82	<0,00001
ARH(1)-CS	2	30,67	<0,00001	ARH(1)-CS	2	28,66	<0,00001
ARH(1)- VC	3	34,00	<0,00001	ARH(1)-VC	3	31,50	<0,00001
ARH(1)-PSC	3	34,00	<0,00001	ARH(1)- PSC	3	31,05	<0,00001
CSH-AR(1)	2	26,73	0,000002	CSH-AR(1)	2	23,31	0,000009
CSH-TOEP	1	26,72	<0,00001	CSH-TOEP	1	23,14	0,000002
CSH-CS	2	29,45	<0,00001	CSH-CS	2	26,98	<0,00001
CSH-VC	3	32,77	<0,00001	CSH-VC	3	29,82	0,000002
CSH-PSC	3	32,77	<0,00001	CSH- PSC	3	29,82	0,000002
HF-AR(1)	2	20,14	0,000042	HF-AR(1)	2	17,20	0,000184
HF-TOEP	1	20,12	0,000007	HF-TOEP	1	17,03	0,000037
HF-CS	2	22,85	0,000011	HF-CS	2	20,86	0,000030
HF-VC	3	26,18	0,000009	HF-VC	3	23,70	0,000029
HF-PSC ²	3	26,18	0,000009	HF- PSC	3	23,70	0,000029
AR(1)-TOEP	1	0,020	0,887537	AR(1)-TOEP	1	0,170	0,680112
AR(1)-VC	1	6,040	0,013985	AR(1)-VC	1	6,510	0,010727
AR(1)-PSC	1	6,040	0,013985	AR(1)-PSC	1	6,510	0,010727
TOEP-CS	1	2,730	0,098479	TOEP-CS	1	3,840	0,050044
TOEP-VC	2	6,060	0,048316	TOEP-VC	2	6,680	0,035437
TOEP-PSC	2	6,060	0,048316	TOEP-PSC	2	6,680	0,035437
CS-VC	1	3,330	0,068027	CS-VC	1	2,840	0,091944
CS-PSC	1	3,330	0,068027	CS- PSC	1	2,840	0,091944

¹Diferença do número de parâmetros entre os dois modelos.

²Para obtenção de valores de AIC comparáveis, o modelo PSC (parcelas subdivididas), correspondente à estrutura de erros independentes, foi ajustado sem o termo relativo ao erro "a" (genótipos x blocos).

estrutura para a matriz Σ . Pela semelhança com o ensaio anterior, no mesmo local, estes resultados não estão aqui apresentados; são disponíveis em SILVA (2011). No caso de genótipos fixos, destacaram-se as estruturas TOEP, CS, CSH e UN; porém, nenhuma destas superou estatisticamente, em nível de 5% de probabilidade, a matriz não estruturada (UN). FREITAS et al. (2005) também destacaram a qualidade de ajuste associada a esse tipo de matriz de covariâncias. Sob a suposição de genótipos aleatórios, outra vez destacaram-se TOEP, UN e CSH, as quais não diferiram entre si ($P>0,05$). Entre as piores

estruturas, novamente estiveram as matrizes VC e PSC, que não se diferiram. Logo, o resultado reforça a inadequação das estruturas de erros independentes para o bom ajustamento dos dados.

Para o ensaio conduzido na localidade de Goianésia, outra vez a suposição associada aos efeitos dos genótipos (fixos ou aleatórios) praticamente não alterou a ordem de classificação das estruturas de Σ (Tabela 4); e, novamente, destacou-se ($P<0,01$), com melhor qualidade de ajustamento, a matriz não estruturada (UN). Também se ratificou o pior desempenho das estruturas PSC e VC em relação

Tabela 3 - Resultados do ajustamento de modelos com diferentes estruturas de matriz de variâncias-covariâncias residuais (Σ), segundo o número de parâmetros (NP), valor máximo da função de verossimilhança (-2 Reml) e do critério de informação de Akaike (AIC), em análises de medidas repetidas de ensaio de competição de variedades de cana-de-açúcar, conduzido em Goiatuba-GO (2004-2008), sob avaliação de quatro cortes e efeitos genotípicos assumidos como fixos ou aleatórios.

Genótipos fixos				Genótipos aleatórios			
Modelos	NP	-2 Reml	AIC	Modelos	NP	-2 Reml	AIC
CSH	5	3065,140	3075,140	CSH	7	3972,133	3986,133
UN	10	3056,470	3076,470	UN	12	3964,526	3988,526
HF	5	3074,578	3084,578	TOEP	6	3981,980	3993,980
ARH(1)	5	3075,518	3085,518	CS	4	3987,477	3995,477
TOEP	4	3077,633	3085,633	HF	7	3982,307	3996,307
CS	2	3083,360	3087,360	ARH(1)	7	3982,403	3996,403
AR(1)	2	3092,690	3096,690	AR(1)	4	3996,816	4004,816
VC	1	3095,808	3097,808	PSC	3	3999,966	4005,966
PSC ¹	1	3095,808	3097,809	VC	3	4000,000	4006,000

¹ Para obtenção de valores de AIC comparáveis, o modelo PSC (parcelas subdivididas), correspondente à estrutura de erros independentes, foi ajustado sem o termo relativo ao erro "a" (genótipos x blocos).

às demais. Ainda vale reportar que a suposição de genótipos fixos também não influenciou nas estimativas das médias genotípicas, mesmo sob as diferentes estruturas de matriz de variâncias-covariâncias (SILVA, 2011). Este resultado corrobora a tese de que a suposição de tratamentos fixos é independente do tipo de estrutura de Σ , se os conjuntos de dados forem balanceados e ortogonais (DUARTE & VENCOVSKY, 2001; CANDIDO, 2009).

CONCLUSÃO

A análise clássica sob erros independentes ($\Sigma=I\sigma^2$), denotada como modelo de parcelas

subdivididas (PSC), oscilou entre as piores estruturas, ou, apenas, entre estruturas medianas quanto à qualidade de ajustamento dos dados. Logo, não se mostra adequada para a análise de ensaios varietais com medidas repetidas em cana-de-açúcar. Além disso, a melhor estrutura quanto à capacidade de o modelo ajustar-se aos dados observados varia entre ensaios, o que revela impossibilidade de se indicar previamente uma estrutura para a análise desse tipo de experimento. Por fim, a suposição relacionada aos efeitos genotípicos (fixos ou aleatórios), necessária na abordagem de modelos mistos, praticamente não altera a classificação das estruturas de matriz de variâncias-covariâncias residuais.

Tabela 4 - Resultados do ajustamento de modelos com diferentes estruturas de matriz de variâncias-covariâncias residuais (Σ), segundo o número de parâmetros (NP), valor máximo da função de verossimilhança (-2 Reml) e do critério de informação de Akaike (AIC), em análises de medidas repetidas de ensaio de competição de variedades de cana-de-açúcar, conduzido em Goianésia-GO (2005-2009), sob avaliação de quatro cortes e efeitos genotípicos assumidos como fixos ou aleatórios.

Genótipos fixos				Genótipos aleatórios			
Modelos	NP	-2 Reml	AIC	Modelos	NP	-2 Reml	AIC
UN	10	1961,266	1981,266	UN	12	2804,886	2828,886
AR(1)	2	1999,396	2003,396	AR(1)	4	2839,574	2847,574
TOEP	4	1995,417	2003,417	TOEP	6	2837,276	2849,276
HF	5	1993,888	2003,888	ARH(1)	7	2838,935	2852,935
ARH(1)	5	1998,233	2008,233	HF	7	2839,959	2853,959
CS	2	2011,828	2015,828	CS	4	2860,306	2868,306
CSH	5	2010,601	2020,601	CSH	7	2855,911	2869,911
PSC ¹	1	2049,100	2051,100	PSC	3	2898,201	2902,201
VC	1	2049,100	2051,100	VC	3	2898,201	2902,201

¹ Para obtenção de valores de AIC comparáveis, o modelo PSC (parcelas subdivididas), correspondente à estrutura de erros independentes, foi ajustado sem o termo relativo ao erro "a" (genótipos x blocos).

AGRADECIMENTOS

Ao Programa de Melhoramento Genético da Cana-de-Açúcar (PMGCA), Universidade Federal de Goiás (UFG), vinculado à Ridesa Brasil, pela seção dos dados e pela bolsa de estudos concedida ao primeiro autor, durante seu curso de mestrado no Programa de Pós-Graduação em Agronomia.

REFERÊNCIAS

- CAMARINHA FILHO, J.A. **Modelos lineares mistos**: estruturas de matrizes de variâncias e covariâncias e seleção de modelo. 2002. 85f. Tese (Doutorado em Agronomia: Estatística e Experimentação Agronômica) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, SP.
- CANDIDO, L.S. **Modelos mistos na avaliação e ordenação de genótipos de cana-de-açúcar, com e sem efeitos de competição com parcelas vizinhas**. 2009. 76f. Tese (Doutorado em Agronomia: Genética e Melhoramento de Plantas) - Faculdade de Ciências Agrárias e Veterinária Campus Jaboticabal, Universidade Estadual Paulista "Júlio de Mesquita Filho", Jaboticabal, SP.
- CECON, P.R. et al. Análise de medidas repetidas na avaliação de clones de café 'Conilon'. **Pesquisa Agropecuária Brasileira**, v.43, p.1171-1176, 2008. Disponível em: <<http://seer.sct.embrapa.br/index.php/pab/article/view/241/5497>>. Acesso em: 17 out. 2014. doi: 10.1590/S0100-204X2008000900011.
- COSTA, S.C. **Modelos lineares generalizados para dados longitudinais**. 2003. 110f. Tese (Doutorado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, SP.
- DUARTE, J.B.; VENCOSKY, R. Estimção e predição por modelo linear misto com ênfase na ordenação de médias de tratamentos genéticos. **Scientia Agricola**, v.58, p.109-117, 2001. Disponível em: <<http://www.scielo.br/pdf/sa/v58n1/a17v58n1.pdf>>. Acesso em: 17 out. 2014. doi: 10.1590/S0103-90162001000100017.
- FLORIANO, E.P. et al. Ajuste e seleção de modelos tradicionais para série temporal de dados de altura de árvores. **Ciência Florestal**, v.16, n.2, p.177-199, 2006. Disponível em: <<http://coral.ufsm.br/cienciaflorestal/artigos/v16n2/A6V16N2.pdf>>. Acesso em: 17 out. 2014.
- FREITAS, A.R. et al. Alternativas de análises em dados de medidas repetidas de bovinos de corte. **Revista Brasileira de Zootecnia**, v.34, p.2233-2244, 2005. Disponível em: <<http://www.scielo.br/pdf/rbz/v34n6s0/a10v3460.pdf>>. Acesso em: 17 out. 2014. doi: 10.1590/S1516-35982005000700010.
- FREITAS, E.G. **Análise de dados longitudinais em experimentos com cana-de-açúcar**. 2007. 75f. Dissertação (Mestrado em Estatística e Experimentação agrônômica) - Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, SP.
- FREITAS, E.G. et al. Modelo univariado aplicado a dados longitudinais de cana-de-açúcar. **Revista Brasileira de Biometria**, v.26, p.93-16, 2008. Disponível em: <http://jaguar.fcav.unesp.br/RME/fasciculos/v26/v26_n2/A6_Edjane.pdf>. Acesso em: 17 out. 2014.
- MAUCHLY, J.W. Significance test for a normal n-variate distribution. **Annals of Mathematical Statistics**, v.11, p.204-209, 1940.
- QUINTAL, S.S.R. **Melhoramento da goiabeira *P. guajava* via metodologia de modelos mistos**. 2013. 193f. Tese (Doutorado em Produção Vegetal) - Programa de Pós-Graduação em Produção Vegetal, Universidade Estadual do Norte Fluminense, RJ. Disponível em: <http://uenf.br/Uenf/Downloads/PRODVEGETAL_3434_1379508554.pdf>. Acesso em: 17 out. 2014.
- RESENDE, M.D.V. et al. Multivariate spatial statistical analysis of longitudinal data in perennial crops. **Revista de Matemática e Estatística**, v.24, p.147-169, 2006. Disponível em: <http://jaguar.fcav.unesp.br/RME/fasciculos/v24/v24_n1/A9_MDLeon.pdf>. Acesso em: 17 out. 2014.
- SILVA, E.N. **Análise de medidas repetidas em ensaios varietais de cana-de-açúcar**. 2011. 96f. Dissertação (Mestrado em Agronomia: Genética e Melhoramento de Plantas) - Programa de Pós-Graduação em Agronomia, Universidade Federal de Goiás, GO. Disponível em: <http://www.files.scire.net.br/atrio/ufg-ppga_upl/THESIS/252/dissertao__emerson_noletto.pdf>. Acesso em: 17 out. 2014.