

DESCRIÇÃO DA VARIABILIDADE INTRA-ESPECÍFICA EM OBSERVAÇÕES DE PESO-IDADE DE BOVINOS POR FUNÇÕES DE COVARIÂNCIA

DESCRIPTION OF INTRA-SPECIFIC VARIABILITY OF AGE-WEIGHT BOVINE RECORDS BY COVARIANCE FUNCTIONS

Santoro, K.R.¹, S.B.P. Barbosa², L.H.A. Brasil³ e E.S. Santos⁴

¹Professor UAG/UFRPE. Rua Ernesto Dourado, 82. Bairro Heliópolis. 55296-190, Garanhuns (PE). Brasil. Autor correspondente. E-mail: ksantoro@ufrpe.br

²Professor DZ/UFRPE. Rua Dom Manoel de Medeiros, s/n. Bairro de Dois Irmãos. 52171-900, Recife (PE). E-mail: sbarbosa@ufrpe.br

³Professora DZ/UFRPE. Rua Dom Manoel de Medeiros, s/n. Bairro de Dois Irmãos. 52171-900, Recife (PE). E-mail: lubrasil@ufrpe.br

⁴Professor DEINFO/UFRPE. Rua Dom Manoel de Medeiros, s/n. Bairro de Dois Irmãos. 52171-900, Recife (PE). E-mail: easant@zaz.com.br

PALAVRAS CHAVE ADICIONAIS

Matriz de covariância. Medidas repetidas. REML. Teste de esfericidade.

ADDITIONAL KEYWORDS

Covariance matrix. REML. Repeated measures. Sphericity test.

RESUMO

A covariância intra-individual geralmente é representada por uma matriz com estrutura esférica, definida por uma função de covariância denominada simetria composta, através de um modelo conhecido como modelo de repetibilidade. Entretanto, esta estrutura pode não ser a mais adequada, principalmente no caso da ocorrência de autocorrelação e heterogeneidade de variâncias. O objetivo deste trabalho foi analisar diferentes estruturas de covariância para modelar a covariância intra-individual e selecionar a mais adequada. Foram utilizados 1737 dados de peso-idade de 193 bovinos Guzerá, nascidos e criados no estado de Pernambuco entre 1970 e 2000, inclusive. Os dados foram analisados por um modelo misto e utilizando doze diferentes estruturas para a matriz de covariâncias intra-individual, sendo ajustados por máxima vero-

ssimilhança restrita. As matrizes mais adequadas foram selecionadas pelo critério de informação de Akaike (AIC) e a melhor predição das correlações entre pesagens. Os efeitos incluídos no modelo tiveram sua significância influenciada pela estrutura da matriz. A matriz que obteve melhor ajuste juntamente a uma boa predição da correlação entre as diferentes pesagens foi a estrutura definida pela função de covariância Markoviana (SP(POW)).

SUMMARY

The intra-individual covariance is usually represented by a spherical structure, defined by covariance function denominated compound symmetry, toward a model named repeatability

Arch. Zootec. 55 (211): 251-262. 2006.

model. This structure, however, could not be the more adequate, particularly in cases in which occur autocorrelation and variance heterogeneity. The aim of this work was to analyze different covariance structures to model the intra-individual covariance and to select the more appropriate. There were utilized 1737 records from 193 animals of Guzerá breed cattle, born and reared in Pernambuco state, northeastern Brazil, between 1970 and 2000, inclusive. The records were analyzed by a mixed model and using twelve different structures for intra-individual covariance matrix, being adjusted by a restricted maximum likelihood process. The more adequate matrix was selected by Akaike information criteria and best prediction of correlations between weightings. The effects included in model had its significance affected by matrix structure. The matrix that had the best adjust together a good prediction of correlations between different weightings was the structure defined by Markovian covariance function (SP(POW)).

INTRODUÇÃO

Observações de medidas repetidas podem ser analisadas de forma univariada ou multivariada, em modelos de efeitos fixos ou mistos (Littell *et al.*, 1996). A análise multivariada não impõe nenhuma restrição ou condição para a matriz de covariâncias dos efeitos, mas na análise univariada a estrutura da matriz deve ser especificada (Khattree e Naik, 1999).

As análises multivariadas ou univariadas geralmente empregam uma abordagem com modelos fatoriais ou então em esquema de parcelas subdivididas, na qual as diferentes observações são representadas pelas subparcelas (Littel *et al.*, 1996; Johnson e Wichern, 1998). Desta forma, as

observações são tomadas como sendo cada uma, uma característica *diferente*, isolada, produzindo um modelo de *janelas* ou *finito* (Kirkpatrick *et al.*, 1990; Pletcher e Geyer, 1999).

Em um modelo de análise de medidas repetidas sob forma univariada em parcelas subdivididas, o teste F para efeitos dentro de indivíduos, ou seja, para observações na mesma unidade experimental, está sujeito a algumas condições para que se verifique a sua validade (Littell *et al.*, 1998). Entre estas condições estão a não heterogeneidade de variâncias entre diferentes observações e o comportamento constante da covariância entre elas (Malheiros, 2001). Uma matriz de covariância que segue este formato ou estrutura é dita como satisfazendo a condição de esfericidade (Huyn e Feldt, 1970; Milliken e Johnson, 1992), a qual pode ser verificada pelo teste de esfericidade de Mauchly (Mauchly, 1940). Entre as matrizes que satisfazem esta condição estão as de estrutura do tipo simetria composta (CS) e a Huyn-Feldt (HF), sendo que a segunda apresenta condições mais abrangentes que a primeira. A estrutura CS é referida como sendo a utilizada por modelos de *repetibilidade* para a variável observada (Meyer, 2000), e é geralmente empregada em análise de medidas repetidas.

As condições citadas geralmente não são satisfeitas em dados de campo (Littel *et al.*, 1991; Davidian e Giltinan, 1995; Littell *et al.*, 1998; Khattree e Naik, 1999). Nestes casos, são sugeridas correções para os graus de liberdade, a fim de corrigir o problema, tais como as de Huynh-Feldt (HF) (Huynh e Feldt, 1976) e Geisser-

Greenhouse (GG) (Geisser e Greenhouse, 1954). Tais correções tornam possível utilizar a distribuição F central de forma aproximada na análise da variável sob repetição (Malheiros, 2001).

Uma outra abordagem em medidas repetidas, além do modelo finito, é a suposição de que a característica estaria mudando continuamente, proporcionando um comportamento de dimensão *infinita* (Kirkpatrick e Heckman, 1989; Jaffrézic e Pletcher, 2000). Entretanto, este modelo *infinitesimal* (Kirkpatrick *et al.*, 1990) impõe a necessidade de uma matriz de covariâncias de dimensão também *infinita*. Uma solução encontrada para esta abordagem está na *função de covariância* (Henderson Jr., 1982; Longford, 1993), que especifica uma função apta a descrever o comportamento dos parâmetros da matriz de covariância em qualquer ponto desejado, inclusive naqueles não observados (Littell *et al.*, 1996). As funções de covariância reduzem o número de parâmetros a serem estimados, diminuindo o esforço computacional e facilitando a compreensão dos resultados encontrados por parte do pesquisador. As funções de covariância podem ser aplicadas tanto para os efeitos chamados *entre* objetos ou *inter-específicos*, quanto para efeitos *dentro* de objetos ou *intra-específicos* (Longford, 1993; Littell *et al.*, 1996).

As funções de covariância têm encontrado emprego crescente nas mais diversas áreas. Em dados espacialmente relacionados como *okriging* na agricultura (Marx e Thompson, 1987), geoestatística (Clark, 1979), estatística espacial (Ripley, 1981), em

dados temporalmente relacionados como experimentos nutricionais (Littel *et al.*, 1998), dados clínicos (Patel, 1986), crescimento animal (Kirkpatrick *et al.*, 1990), melhoramento genético animal (Meyer e Hill, 1997; Meyer, 1999; Meyer, 2000), regressão aleatória (Meyer, 1998), entre outras (Diggle *et al.*, 1994; Lindsey, 1993). Isto se deve principalmente à eficiente implementação da metodologia em pacotes estatísticos de análise de dados, tais como o SAS (SAS Institute, 2000) e S-PLUS (MATHSOFT, 1999).

Entre as diversas configurações possíveis de serem atribuídas à matriz de covariância através de funções de covariância, nem todas seriam apropriadas. A escolha da função de covariância mais adequada seria um processo interativo de análise, que integraria o objetivo do trabalho, a amostra em estudo, o modelo estatístico empregado e o ajuste oferecido pelas diferentes funções de covariância possíveis de serem aplicados (Littel *et al.*, 1998; Keselman *et al.*, 1998). Uma das conseqüências da escolha da matriz é que os testes estatísticos são afetados por sua configuração (Wolfinger e Chang, 1995; Malheiros, 2001).

Observações de crescimento, provenientes de experimentos controlados ou não, oferecem um interessante campo de aplicação, pois os dados encontram-se influenciados por efeitos fixos e aleatórios, proporcionando uma abordagem por modelo misto. Apresentam ainda, heterocedasticidade e correlação entre e dentro das unidades experimentais, as quais geralmente são representadas pelos indivíduos (Davidian e Giltinan, 1995; Núñez-Antón e Woodworth, 1994). Além disso, como

os dados são temporalmente relacionados, há um comportamento de que medidas em tempos distantes sejam menos correlacionadas do que aquelas em tempos próximos (Littell *et al.*, 1998).

Este trabalho teve por objetivo a identificação de uma função de covariância mais adequada para descrever efeitos dentro de unidades experimentais em dados de crescimento do tipo peso-idade, de bovinos Guzerá criados no estado de Pernambuco.

MATERIAL E MÉTODOS

A partir de 193 animais Guzerá, nascidos entre 1960 e 2000 e criados no estado de Pernambuco, selecionou-se animais pelo critério de necessariamente possuírem todas as pesagens (nove), desde o nascimento, até dois anos de idade, sendo que essas pesagens ocorreram, aproximadamente, a cada noventa dias, o que totalizou 1737 observações de peso-idade. Os animais eram filhos de 54 pais e 173 mães. Os animais constavam do banco de dados de Controle de Desenvolvimento Ponderal da Associação Brasileira de Criadores de Zebu (ABCZ).

Os animais apresentaram diferentes sistemas de criação, a saber: pasto, estabulado e com mudança de um sistema de criação para outro.

Utilizou-se o seguinte modelo misto para a descrição dos dados de pesagem:

$$Y_{ijklmnop} = \mu + S_k + P_n + T_o + A_l + M_m + D + D^2 + D^3 + R_q + F_r + e_{ijklmnop}$$

onde: $Y_{ijklmnop}$ representa o peso observado para

o animal i na pesagem j , $i = 1, \dots, 193$; $j = 1, \dots, 9$; μ representa uma constante comum a todas as observações; S_k representa o efeito do sexo k ($1 = \text{fêmea}$, $2 = \text{macho}$); P_n representa o efeito da propriedade n ($n = 1, \dots, 11$); T_o representa o efeito do regime de criação o ($1 = \text{pasto}$, $2 = \text{estabulado}$, $3 = \text{com mudança}$); A_l representa o ano de nascimento l ($l = 1980, \dots, 1996$); M_m representa o efeito do mês de nascimento m ($m = 1, \dots, 12$); D , D^2 , D^3 representam os efeitos linear, quadrático e cúbico, respectivamente, do fator tempo (dias) desde o nascimento do animal i até a pesagem j ; R_q representa o efeito aleatório do reprodutor q ($q = 1, \dots, 54$); F_r representa o efeito aleatório da mãe r ($r = 1, \dots, 173$); $e_{ijklmnop}$ representa o erro aleatório associado a cada observação.

O modelo misto pode ser escrito matricialmente como:

$$y = Xb + Za + e$$

onde y denota o vetor de observações, X é a matriz de efeitos fixos conhecidos, b é o vetor de parâmetros de efeitos fixos desconhecidos, Z é a matriz de efeitos aleatórios conhecidos, a é o vetor de parâmetros de efeitos aleatórios desconhecidos, e é o vetor não observado de erros aleatórios independentemente, identicamente e normalmente distribuídos. Assume-se que

$$E \begin{bmatrix} a \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ e } \text{Var} \begin{bmatrix} a \\ e \end{bmatrix} = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix}$$

A variância de y seria, então, $V = ZGZ' + R$. Pode-se modelar V através da matriz de efeitos aleatórios Z e pela especificação de estruturas de covariância para G e R . A matriz R é geralmente estruturada como $R = \sigma^2 I_n$,

onde I_n é uma matriz identidade $n \times n$.

VARIABILIDADE INTRA-ESPECÍFICA EM OBSERVAÇÕES DE PESO-IDADE DE BOVINOS

Para a matriz de covariâncias dos efeitos dentro das pesagens (intra-individuais) R foram ajustadas as matrizes com estrutura simetria composta (CS), simetria composta heterogênea (CSH), autorregressiva de ordem um (AR(1)), autorregressiva de médias móveis de ordem um (ARMA(1,1)), Toeplitz (TOEP), Toeplitz heterogênea (TOEPH), componente de variância (VC), Markoviano ou *spatial power* (SP(POW)), Huyn-Feldt (HF), ante-dependência de ordem um (ANTE(1)), fator analítico de ordem um (FA(1)) e fator analítico de ordem um com diagonal igual (FA1(1)). Ajustou-se uma matriz não estruturada com covariâncias para poder se realizar o teste de esfericidade de Mauchly (Mauchly, 1940). As estruturas, números de parâmetros a estimar (**tabela I**) estão conforme Wolfinger (1993), Macchiavelli e Arnold (1994), Wolfinger (1996) e SAS Institute (2000).

Os parâmetros do modelo misto foram estimados por máxima verossimilhança restrita, através do PROC MIXED do programa SAS (SAS Institute, 2000).

A possibilidade de diferença significativa entre o ajuste de modelos com diferentes matrizes foi realizada através do teste de razão de verossimilhança, descrita por Matsushita (1994), Khattree e Naik (1999) e Xavier (2000), quando possível.

Os seguintes critérios de ajuste estavam disponíveis (Wolfinger, 1993):

- a) $-2 \ln l = -2 \log$ verossimilhança:
- b) Critério de informação de Akaike (AIC) = $-2 \ln l + 2q$
- c) Critério Bayesiano de Schwartz

$$(BIC) = -2 \ln l + q \log (n - p)$$

onde q é o número de parâmetros de covariância, n é o número de observações e p é o posto de X .

A matriz de melhor ajuste foi escolhida pelo melhor valor de ajuste do critério de informação de Akaike (AIC) juntamente a seu comportamento na predição das covariâncias e correlações apresentadas na amostra.

RESULTADOS E DISCUSSÃO

O teste de esfericidade de Mauchly realizado, conforme descrito por Matsushita (1994), Khattree e Naik (1999) e Malheiros (2001), utilizou os dados dos ajustes apresentados pelas estruturas HF e UM (**tabela II**) e demonstrou a não esfericidade da matriz de covariâncias (g.l. = 35; $\chi^2 = 1017,6$; $p < 0,01$). Desta forma, outras estruturas para a matriz de foram testadas.

Entre os critérios de seleção para o melhor modelo misto estão o teste assintótico da razão de verossimilhança, o critério de informação de Akaike (AIC) e o critério bayesiano de Schwarz (BIC). O primeiro, entretanto, só pode ser utilizado para comparar um modelo contra outro, desde que uma matriz de covariância seja um caso especial da outra (Matsushita, 1994). Havendo várias matrizes a serem comparadas, e algumas não sendo casos especiais de outras, não foi permitida a utilização do teste assintótico da razão de verossimilhança para comparação entre todas elas; escolheu-se, então, a melhor matriz pelo critério AIC. O AIC pode

Tabela I. Nome da estrutura, números de parâmetros a estimar e configuração dos elementos da estrutura. (Structure name, number of parameters to be estimated, and elements structure configuration).

Estrutura	Número de parâmetros de covariância	(i, j) -ésimo elemento
ANTE(1)	$2t - 1$	$\mathbf{s}_i \mathbf{s}_j \prod_{k=i}^{j-1} \mathbf{r}_k$
AR(1)	2	$\mathbf{s}^2 \mathbf{r}^{ i-j }$
ARMA(1,1)	3	$\mathbf{s}^2 [\mathbf{g} \mathbf{r}^{ i-j -1} \mathbf{1}(i \neq j) + \mathbf{1}(i = j)]$
CS	2	$\mathbf{s}_i^2 + \mathbf{s}^2 \mathbf{1}(i = j)$
CSH	$t + 1$	$\mathbf{s}_i \mathbf{s}_j [\mathbf{r} \mathbf{1}(1 \neq j) + \mathbf{1}(i = j)]$
FA1(q)	$[q/2](2t - q + 1) + 1$	$\sum_{k=1}^{\min(i, j, q)} \mathbf{l}_{ik} \mathbf{l}_{jk} + \mathbf{s}^2 \mathbf{1}(i = j)$
FA(q)	$[q/2](2t - q + 1) + t$	$\sum_{k=1}^{\min(i, j, q)} \mathbf{l}_{ik} \mathbf{l}_{jk} + \mathbf{s}_i^2 \mathbf{1}(i = j)$
HF	$t + 1$	$(\mathbf{s}_i^2 + \mathbf{s}_j^2) / 2 + \mathbf{1} \mathbf{1}(i \neq j)$
SP(POW)	2	$\mathbf{s}^2 \mathbf{r}^{d_{ij}}$
TOEP	t	$\mathbf{s}_{ i-j +1}$
TOEPH	$2t - 1$	$\mathbf{s}_i \mathbf{s}_j \mathbf{r}_{ i-j }$
UN	$t(t + 1) / 2$	\mathbf{s}_{ij}
VC	q	$\mathbf{s}_k^2 \mathbf{1}(i = j)$

ρ é o parâmetro autoregressivo, sendo $|\rho| < 1$ para estacionaridade; γ modela o componente de médias móveis, sendo $|\gamma| < 1$ para estacionaridade; d_{ij} é a distância Euclidiana entre os i -ésimo e j -ésimo vetores das coordenadas, que correspondem as i -ésima e j -ésima observações; t é a dimensão total da matriz de covariância; q é o número fatores. (ρ is the autoregressive parameter, and $|\rho| < 1$ for stationarity; γ models the moving average component, and $|\gamma| < 1$ for stationarity; d_{ij} is the Euclidian distance between the i -th and j -th coordinate vectors, that corresponds to i -th and j -th observations; t is the overall dimension of the covariance matrix; q is the number of factors).

VARIABILIDADE INTRA-ESPECÍFICA EM OBSERVAÇÕES DE PESO-IDADE DE BOVINOS

ser utilizado para comparar modelos com os mesmos efeitos fixos, mas diferentes estruturas de variância, sendo que o modelo com o menor AIC será considerado o melhor. No caso do BIC o julgamento é o mesmo, mas ele penaliza mais os modelos com grande número de parâmetros do que o AIC faz, sendo que os dois critérios podem não concordar (Wolfinger e Chang, 1995; Xavier, 2000).

A matriz mais adequada segundo o critério AIC para a descrição dos efeitos intra-específicos foi a do tipo FA(1), seguida pela SP(POW) (**tabela II**).

O teste de razão de verossimilhança entre todos os ajustes oferecidos pelas diferentes estruturas de covariância estudadas (**tabela III**) demonstrou que somente as estruturas FA1(1) e ARMA(1,1) e as estruturas SP(POW) e CSH ofereceram igual ajuste. Algumas estruturas, entretanto, não

possibilitaram a realização do teste devido a terem o mesmo número de graus de liberdade: CS-AR(1), SP(POW)-AR(1), VC-ARMA(1,1), SP(POW)-CS, HF-CSH. A diferença entre as estruturas HF e ARMA(1,1) foi significativa a 5 p.100 ($p < 0,05$), sendo que todas as demais foram significativas a 1 p.100 ($p < 0,01$).

As diferentes estruturas resultaram em diferentes significâncias para os efeitos fixos estudados no modelo (**tabela IV**). Ambos comportamentos seriam esperados, conforme comentado por Wolfinger e Chang (1995), uma vez que os testes estatísticos seriam uma função da estrutura utilizada, sendo que resultados compatíveis foram obtidos por outros autores (Wolfinger, 1993; Wolfinger, 1996; Littell *et al.*, 1998; Malheiros, 2001).

Os efeitos com maior número de ocorrências de significância para as diferentes estruturas foram (**tabela**

Tabela II. Ajuste obtido com o uso das diferentes estruturas da matriz de covariâncias para a descrição dos efeitos intra-específicos. (Obtained adjust by different structures of covariance matrix for description of intra-specific effects).

Estrutura	Parâmetros	-2 log verossimilhança	AIC	BIC
ANTE(1)	16	79736,3	79768,3	79799,5
AR(1)	2	16324,4	16328,4	16332,4
ARMA(1,1)	3	16291,5	16297,5	16303,4
CS	2	17976,9	17980,9	17984,8
CSH	12	16135,2	16159,2	16183,1
FA1(1)	11	16286,9	16308,9	16330,8
FA(1)	20	15800,1	15840,1	15879,9
HF	12	16271,3	16295,3	16318,8
SP(POW)	2	16131,9	16135,9	16139,9
TOEP	9	16189,3	16207,3	16225,2
TOEPH	15	72472,3	72502,3	72532,2
UN	47	15253,7	15347,7	15441,2
VC	3	17995,5	18001,5	18007,5

Tabela III. Teste de razão de verossimilhança para os ajustes obtidos pelas diferentes estruturas da matriz de covariâncias. (Likelihood ratio test for adjust obtained by different covariance matrix structures).

Estrutura	-2 log ver. ^a	ANTE(1)	AR(1)	ARMA(1,1)	CS	CSH	FA1(1)	FA(1)	HF	SP(POW)	TOEP	TOEPH	VC
ANTE(1)	79736,3												
	(16)												
AR(1)	16324,4	63411,9											
	(2)	(14)											
ARMA(1,1)	16291,5	63444,8	32,9										
	(3)	(13)	(1)										
CS	17976,9	61759,4	1652,5^b	1685,4									
	(2)	(14)	(0)	(1)									
CSH	16135,2	63601,1	189,2	156,3	1841,7								
	(12)	(4)	(10)	(9)	(10)								
FA1(1)	16286,9	63449,4	37,5	4,6 ^{NS}	1690,0	151,7							
	(11)	(5)	(9)	(8)	(9)	(1)							
FA(1)	15800,1	63936,2	524,3	491,4	2176,8	335,1	486,8						
	(20)	(4)	(18)	(17)	(18)	(8)	(9)						
HF	16271,3	63465,0	53,1	20,2 [*]	1705,6	136,1^b	15,6	471,2					
	(12)	(4)	(10)	(9)	(10)	(0)	(1)	(8)					
SP(POW)	16131,9	63604,4	192,5^b	159,6	1845,0^b	3,3 ^{NS}	155,0	331,8	139,4				
	(2)	(14)	(0)	(1)	(0)	(10)	(9)	(18)	(10)				
TOEP	16189,3	63547,0	135,1	102,2	1787,6	54,1	97,6	389,2	82,0	57,4			
	(9)	(7)	(7)	(6)	(7)	(3)	(2)	(11)	(3)	(7)			
TOEPH	72472,3	7264	56147,9	56180,8	54495,4	56337,1	56185,4	56672,2	56201,0	56340,4	56283,0		
	(15)	(1)	(13)	(12)	(13)	(3)	(4)	(5)	(3)	(13)	(6)		
VC	17995,5	61740,8	1671,1	1704,0^b	18,6	1860,3	1708,6	2195,4	1724,2	1863,6	1806,2	54476,8	
	(3)	(13)	(1)	(0)	(1)	(9)	(8)	(17)	(9)	(1)	(6)	(12)	

^aGraus de liberdade entre parênteses; ^bnão foi possível realizar o teste; ^{NS}não significativo; *significativo a 5 p.100 (p<0,05). Todas as demais comparações foram significativas (p<0,01), exceto quando indicado.

^aDegrees of freedom between braces; ^bit was not possible to do the test; ^{NS}non significant; *significant at 5 percent (p<0,05). All others comparisons were significant (p<0,01), exceptions indicated.

VARIABILIDADE INTRA-ESPECÍFICA EM OBSERVAÇÕES DE PESO-IDADE DE BOVINOS

IV), em ordem, os efeitos linear, quadrático e cúbico do tempo, o ano de nascimento, o regime de criação, a propriedade e o sexo. O efeito de menor influência foi o do mês de nascimento, sendo significativo apenas para três estruturas.

Apesar da estrutura FA(1) ter apresentado melhor ajuste pelo critério AIC que a estrutura SP(POW) (tabela II), a segunda estrutura obteve uma predição mais aproximada das correlações entre pesagens que a FA(1), demonstrando maior adequação aos dados. Este comportamento tornou-se mais evidente ao se analisar o gráfico da figura 1, o qual demonstrou que para os diferentes intervalos (lags) de pesagens, a predição da correlação

pela estrutura SP(POW) esteve mais próxima do apresentado pela amostra que a estrutura FA(1), principalmente para as pesagens distanciadas em um, dois ou oito intervalos. Isto tem implicação direta na utilização do modelo, pois pode levar a erros de predição caso o peso a ser predito caia em um destes intervalos. Por exemplo, ao se predizer o peso de um animal dois intervalos adiante, tomando-se como base o peso atual, a estrutura SP(POW) o fará com menos erro que a FA(1), e o erro aumentará ainda mais se o intervalo for igual a um ou oito (figura 1).

As estruturas SP(POW) e CSH não foram estatisticamente diferentes pelo teste de razão de verossimilhança (tabela III), mas a SP(POW) foi

Tabela IV. Influência de efeitos fixos segundo as diferentes estruturas da matriz de covariâncias. (Fixed effects influence due to different covariance matrix structures).

Estrutura	GL num. GL den.	Efeito							
		Pro 10	S 1	RC 2	AN 12	MN 11	D 1	D ² 1	D ³ 1
ANTE(1)		2,87**	11,54**	28,65**	4,5**	8,66**	1700,88**	827,91**	858,28**
AR(1)		2,39**	6,91**	15,88**	2,02*	0,68 ^{NS}	775,46**	87,37**	50,48**
ARMA(1,1)		2,68**	7,6**	17,81**	2,26**	0,76 ^{NS}	667,37**	72,2**	40,95**
CS		5,43**	8,81**	38,15**	4,87**	1,27 ^{NS}	422,46**	52,02**	32,9**
CSH		1,58 ^{NS}	0,51 ^{NS}	0,12 ^{NS}	2,78**	1,86*	1228,49**	166,79**	75,41**
FA1(1)		3,07**	0,8 ^{NS}	4,9**	6,26**	1,04 ^{NS}	1481,6**	147,11**	83,17**
FA(1)		1,77 ^{NS}	1,47 ^{NS}	3,22*	0,72 ^{NS}	1,23 ^{NS}	2038,71**	261,54**	129,8**
HF		2,36*	2,09 ^{NS}	5,18*	2,37**	1,83*	432,46**	56,28**	35,95**
SP(POW)		2,36**	6,89**	16,24**	2,04*	0,68 ^{NS}	881,75**	106,6**	62,07**
TOEP		4,64**	17,92**	28,53**	3,87**	1,54 ^{NS}	730,38**	91,03**	54,3**
TOEPH		3,01**	11,22**	15,04**	5,4**	9,42**	1002,17**	466,03**	485,93**
VC		4,22**	7,99**	35,21**	3,91**	3**	416,41**	51,05**	32,21**

Pro: Propriedade; S: Sexo; RC: Regime de criação; AN: Ano de nascimento; MN: Mês de nascimento.

*Significativo a 5 p.100 (p<0,05); **significativo a 1 p.100 (p<0,01); ^{NS}não significativo.

*Significative at 5 percent (p<0,05); **significative at 1 percent (p<0,01); ^{NS}non significative.

melhor que a CSH pelo critério AIC (**tabela II**), e entre intervalos de pesagem (**figura 1**). A não ser pelo modelo CSH todos os outros demonstraram correlação decrescente com o aumento da distância entre medidas, o que concorda com os valores observados. Assim, a modificação dos pesos em um programa de melhoramento genético teria mais resultados para pesos mais próximos que para mais distantes. Além disso, pesagens próximas precisariam de menor número de observações que as distantes.

O comportamento da predição das correlações entre pesagens e entre intervalos de pesagem e a simplicidade de descrição e ajuste da estrutura SP(POW) em relação à FA(1), pois são dois parâmetros contra vinte, colocou a estrutura SP(POW) como a mais apropriada, apesar do ajuste superior da FA(1) pelo critério AIC, con-

forme a **tabela II**. Khattree e Naik (1999) comentaram que esta estrutura seria a mais indicada para medidas correlacionadas tomadas em espaços de tempo desiguais.

CONCLUSÕES

A suposição de existência de esfericidade poderia levar o pesquisador a erros de inferência e estimação. Assim, recomenda-se a verificação dessa propriedade através de testes específicos. Tanto efeitos inter quanto intra-específicos são influentes sobre o comportamento das características longitudinais, devendo ser considerados em sua análise e explicação.

A seleção da configuração da matriz de covariâncias para o acondicionamento das relações entre efeitos e dentro de observações deve ser

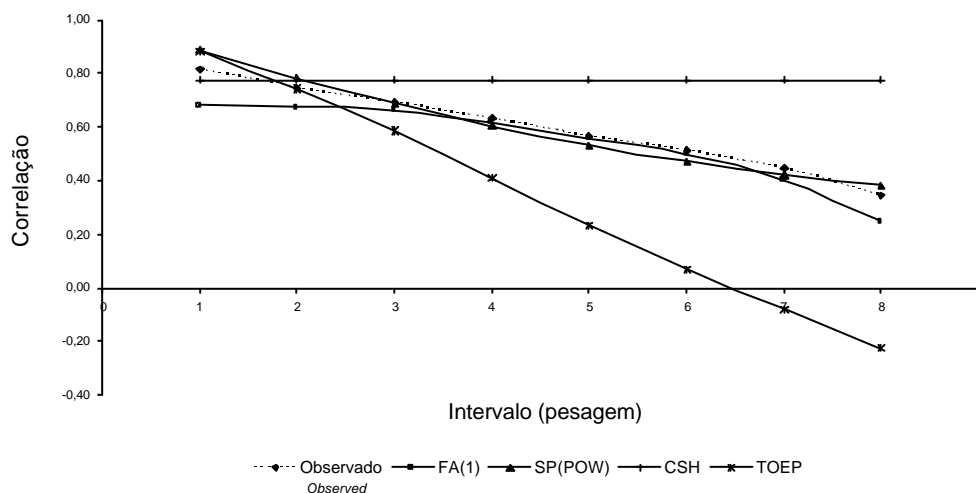


Figura 1. Correlação entre intervalos de pesagens observado na amostra e preditos pelas estruturas FA(1), SP(POW), CSH e TOEP. (Observed correlations in sample and predicted by FA(1), SP(POW), CSH, and TOEP structures).

VARIABILIDADE INTRA-ESPECÍFICA EM OBSERVAÇÕES DE PESO-IDADE DE BOVINOS

baseada em critérios rigorosos e no conhecimento do pesquisador a respeito do comportamento da característica.

Dados de crescimento de bovinos

da raça Guzerá demonstraram ter alta correlação entre medidas mais próximas e menores entre as mais distantes, e a matriz mais adequada foi a de estrutura do tipo SP(POW).

BIBLIOGRAFIA

- Clark, I. 1979. Practical geostatistics. Applied Science Publications. Essex.
- Davidian, M. and D.M. Giltinan. 1995. Nonlinear models for repeated measurement data. Chapman & Hall, London.
- Diggle, P.G., K.Y. Liang and S.L. Zeger. 1994. Analysis of longitudinal data. Clarendon Press, Oxford.
- Geisser, S. and S.W. Greenhouse. 1954. An extension of Box's results on the use of the F distribution in multivariate analysis. *Ann. Math. Statistics*, 29: 484-498.
- Henderson Jr., C.R. 1982. Analysis of covariance in a mixed model: higher-level, nonhomogeneous and random regression. *Biometrics*, 38: 633-640.
- Huynh, H. and L.S. Feldt. 1970. Conditions under which mean square ratios in repeated measurements designs have exact F distributions. *J. Am. Statistics Association*, 65: 1582-1589.
- Huynh, H. and L.S. Feldt. 1976. Estimation of the Box correction for degrees of freedom from sample data in the randomized block and split-plot designs. *J. Educ. Association*, 1: 69-82.
- Jaffrézic, F. and S.D. Pletcher. 2000. Statistical models for estimating the genetic basis of repeated measures and other function-valued traits. *Genetics*, 156: 913-922.
- Johnson, R.A. and D.W. Wichern. 1998. Applied multivariate statistical analysis. 4.ed. Prentice Hall Inc., Englewood Cliffs.
- Keselman, H.J., J. Algina, R.K. Kowalchuck and R.D. Wolfinger. 1998. A comparison of two approaches for selection covariance structures in the analysis of repeated measurements. *Communications in Statistics, Series B*, 27: 591-604.
- Khattree, R. and D.N. Naik. 1999. Applied multivariate statistics with SAS software. 2.ed. SAS Institute Inc., Cary.
- Kirkpatrick, M., D. Lofsvold and M. Bulmer. 1990. Analysis of inheritance, selection and evolution of growth trajectories. *Genetics*, 124: 979-993.
- Kirkpatrick, M. and N. Heckan. 1989. A quantitative genetic model for growth, shape, and other infinite-dimensional characters. *J. Math. Biol.*, 27: 429-450.
- Lindsey, J.K. 1993. Models for repeated measures. Clarendon Press, Oxford.
- Littell, R.C., R.J. Freund e P.C. Spector. 1991. SAS System for linear models. 3a. ed. SAS Institute Inc., Cary (NC).
- Littell, R.C., G.A. Milliken, W.W. Stroup and R.D. Wolfinger. 1996. SAS System for mixed models. SAS Institute Inc., Cary (NC).
- Littell, R.C., P.R. Henry and C.B. Ammerman. 1998. Statistical analysis of repeated measures data using SAS procedures. *J. Anim. Sci.*, 76: 1216-1231.
- Longford, N.T. 1993. Random coefficients models. Clarendon Press, Oxford.
- Macchiaveli, R.E. and S.F. Arnold. 1994. Variable order ante-dependence models. *Communications in Statistics-Theory Methodology*, 23: 2683-2699.
- Malheiros, E.B. 2001. Precisão da análise de experimentos com medidas repetidas usando procedimentos do SAS. *Rev. Matemática e Estatística*, 19: 253-272.

Archivos de zootecnia vol. 55, núm. 211, p. 261.

SANTORO, BARBOSA, BRASIL E SANTOS

- Marx, D. and K. Thompson. 1987. Practical aspects of agricultural Kriging. Arkansas Agricultural Experiment Station, Fayetteville. (Bulletin 903)
- MATHSOFT. 1999. S-PLUS 2000 User's guide. v.1. MathSoft - Data Analysis Products Division, Seattle.
- Matsushita, R.Y. 1994. Modelos longitudinais mistos com correlação serial nos erros. Dissertação de Mestrado em Estatística, Universidade Estadual de Campinas (UNICAMP), Campinas (Brasil).
- Mauchly, J.W. 1940. Significance test for sphericity of a normal n-variate distribution. *Ann. Math. Statistics*, 11: 204-209.
- Meyer, K. 1998. Estimating covariance functions for longitudinal data using a random regression model. *Genet. Sel. Evol.*, 30: 221-240.
- Meyer, K. 1999. Estimates of genetic and phenotypic covariance functions for postweaning growth and mature weight of beef cows. *J. Anim. Breed. Genet.*, 116: 181-205.
- Meyer, K. 2000. Random regressions to model phenotypic variation in monthly weights of Australian beef cows. *Livest. Prod. Sci.*, 65: 19-38.
- Meyer, K. and W.G. Hill. 1997. Estimation of genetic and phenotypic covariance functions for longitudinal data by restricted maximum likelihood. *Livest. Produ. Sci.*, 47: 185-200.
- Milliken, G.A. and D.E. Johnson. 1992. Analysis of messy data. v.1. Designed experiments. Chapman & Hall, New York.
- Núñez-Antón, V. and G.G. Woodworth. 1994. Analysis of longitudinal data with unequally spaced observations and time-dependent correlated errors. *Biometrics*, 50: 445-456.
- Patel, H.I. 1986. Analysis of repeated measures designs with changing covariates in clinical trials. *Biometrika*, 73: 707-715.
- Pletcher, D. and C.J. Geyer. 1999. The genetic analysis of age-dependent traits: modeling the character process. *Genetics*, 151: 825-835.
- Ripley, B.C. 1981. Spatial statistics. John Wiley & Sons, New York.
- SAS Institute. 2000. SAS/STAT User's guide. Version 8. v.2. SAS Institute Inc., Cary (NC).
- Wolfinger, R. 1993. Covariance structure selection in general mixed models. *Communications in Statistics - Simulation*, 22: 1079-1106.
- Wolfinger, R.D. 1996. Heterogeneous variance-covariance structures for repeated measures. *J. Agric. Biol. Envir. S.*, 1: 205-230.
- Wolfinger, R. and M. Chang. 1995. Comparing the SAS GLM and MIXED procedures for repeated measures. In: Proceedings of the Twentieth Annual SAS Users Group Conference, Cary (NC), United States of America.
- Xavier, L.H. 2000. Modelos univariado e multivariado para análise de medidas repetidas e verificação da acurácia do modelo univariado por meio de simulação. Dissertação de Mestrado em Estatística Experimental. Escola Superior de Agricultura "Luiz de Queiroz" (ESALQ)/Universidade de São Paulo (USP), Piracicaba (Brasil).

Recibido: 18-5-05. Aceptado: 21-3-06.

Archivos de zootecnia vol. 55, núm. 211, p. 262.