

INFLUÊNCIA DA ESTRUTURA DA MATRIZ DE COVARIÂNCIAS NA CLASSIFICAÇÃO DE REPRODUTORES CAPRINOS

INFLUENCE OF COVARIANCE MATRIX STRUCTURE TYPE ON MALE GOAT RANKING

Santoro, K.R.¹ e Barbosa, S.B.P.²

¹Unidade Acadêmica de Garanhuns. Universidade Federal Rural de Pernambuco (UFRPE). Av. Bom Pastor, s/n. Bairro Boa Vista. Caixa Postal 152. CEP 55296-901. Garanhuns. Brasil. krsantoro@uag.ufrpe.br

²Departamento de Zootecnia. Universidade Federal Rural de Pernambuco (UFRPE). Rua Dom Manoel de Medeiros, s/n. Bairro de Dois Irmãos. 52171-900 Recife. Brasil. sbarbosa@dz.ufrpe.br

PALAVRAS CHAVE ADICIONAIS

Caprinos Moxotó. Funções de covariância. Medidas repetidas. REML. Teste de esfericidade.

ADDITIONAL KEYWORDS

Covariance functions. Moxotó breed. REML. Repeated measures. Sphericity test.

RESUMO

Analisaram-se dados de peso-idade de 412 caprinos da raça Moxotó, filhos de 151 mães e 50 pais, nascidos entre 1984 e 1988, nos meses de março a julho, no estado de Pernambuco, totalizando 1749 observações. Considerou-se um modelo misto com os seguintes fatores: sexo (S), ano de nascimento (A), mês de nascimento (M), tipo de parto (T), tipo de pelagem (P), ordem de parto (O), efeito linear da umidade relativa (U), efeito linear da temperatura (C), efeito linear da pluviosidade (H), efeito aleatório do reprodutor (R), efeito aleatório da mãe (F), efeito linear (D) e quadrático (D2) do fator tempo em dias. O objetivo do trabalho foi verificar a influência do tipo de matriz de covariância entre pesos sobre a predição de efeitos aleatórios e sua influência na classificação de reprodutores. Ajustaram-se as matrizes simetria composta (CS), simetria composta heterogênea (CSH), autorregressiva de ordem um (AR(1)), autorregressiva heterogênea de ordem um (ARH(1)), autorregressiva de médias móveis de ordem um (ARMA(1,1)), Toeplitz heterogênea (TOEPH), não estruturada com correlações (UNR), componente de variância (VC) e *spatial power* (SP(POW)). A alteração do tipo da matriz de covariância alterou a predição dos efeitos aleatórios e fixos estudados no modelo. A matriz mais adequada foi a ARH(1), os efeitos influentes para o melhor modelo foram reprodutor ($p<0,05$), fêmea ($p<0,01$), sexo ($p<0,01$) e tipo de parto

($p<0,01$) e efeito linear e quadrático do tempo ($p<0,01$).

SUMMARY

Were analysed records of weight-age from 412 goats from Moxotó breed, originated from 151 females and 50 males, born between 1984 and 1988, from march to july, in Pernambuco state, northeastern Brazil, totalizing 1749 records. Was considered a mixed model with the factors: sex (S), year of birth (A), month of birth (M), parturition type (T), bristle type (P), parturition order (O), linear effect of relative umidity (U), linear effect of temperature (C), linear effect of pluviosity (H), random effect of male (R), random effect of female (F), linear (D) and quadratic (D2) time effect in days. The objective was to verify the influence of covariance matrix type between weights on prediction of random effects and the ranking of male goats. Were utilized the follow matrix type compound simmetry (CS), heterogeneity compound simmetry (CSH), autoregressive of order one (AR(1)), heterogeneous autoregressive of order one (ARH(1)), autoregressive moving average of order one (ARMA(1,1)), heterogeneous Toeplitz (TOEPH), non structured with correlation (UNR), variance component (VC), and spatial power (SP(POW)). Change in covariance matrix type changed the predicted random model effects. The

Recibido: 14-2-08. Aceptado: 17-11-08.

Arch. Zootec. 59 (227): 345-355. 2010.

most adequate matrix was ARH(1), influential effects were male ($p < 0,05$), female ($p < 0,01$), sex ($p < 0,01$), parturition type ($p < 0,01$), and linear and quadratic effect of time ($p < 0,01$).

INTRODUÇÃO

A variabilidade existente em dados longitudinais de observações de crescimento animal pode ser devida a fatores entre animais (inter-específicos) e dentro dos próprios animais (intra-específicos) (Khattree e Naik, 1999; Littell *et al.*, 1996).

A diferenciação entre os animais, para posterior julgamento e seleção, visando um programa de melhoramento genético animal, baseia-se geralmente na predição dos valores genéticos dos indivíduos através de avaliações inter-específicas.

A variabilidade intra-específica é, se não esquecida, em muitas vezes assumida como tendo uma estrutura de simetria composta, a qual pode ser descrita por uma matriz de relacionamento (covariâncias ou correlações) entre as diferentes observações que considera a não heterogeneidade de variâncias entre diferentes observações e o comportamento constante da covariância entre elas (Malheiros, 2001).

Esta suposição não satisfaz algumas características comuns em observações de crescimento, pois elas apresentam correlação entre as diferentes medidas, sendo as medidas próximas mais correlacionadas que as distantes e não raro também contam com heterogeneidade de variância (Davidian e Giltinan, 1995; Khattree e Naik, 1999; Littell *et al.*, 1991; Littell *et al.*, 1998; Nunez-Anton e Woodworth, 1994).

A verificação da adequação de uma matriz da classe esférica (Huynh e Feldt, 1970; Milliken e Johnson, 1992), da qual a estrutura simetria composta faz parte, pode ser feita através do teste desenvolvido por Maucly (Mauchly, 1940). Caso a esfericidade não seja comprovada seria necessário testar a adequação de outras estruturas matriciais.

A escolha da forma da matriz mais adequada seria um processo iterativo 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 (Keselman *et al.*, 1998; Littell *et al.*, 1998).

A importância da seleção desta estrutura de relacionamento teria consequências diretas sobre a estimação dos efeitos fixos do modelo e suas significâncias, assim como alteraria as predições dos efeitos aleatórios (Littell *et al.*, 1998; Malheiros, 2001; Wolfinger e Chang, 1995).

Isto poderia ser notado, por exemplo, através da mudança na classificação ou *rank* dos animais. Tal mudança de posição levaria a seleção de diferentes animais como os mais adequados ou com maior potencial genético para serem utilizados em um programa de melhoramento genético animal, o que poderia comprometer o andamento do programa ou seus objetivos. Desta forma, não somente o relacionamento entre os animais e os efeitos comuns entre eles (inter-específicos) são importantes, mas também os que delineiam o comportamento do indivíduo em si (intra-específicos).

A suposição de que uma característica com observações longitudinais poderia ser entendida como estando sob mudança contínua, proporcionaria a ela um comportamento de dimensão *infinita* (Jaffrézic e Pletcher, 2000; Kirkpatrick e Heckman, 1989). Tal representação seria feita por um modelo *infinitesimal* (Kirkpatrick *et al.*, 1990), e teria 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), a qual especifica uma função apta a descrever o comportamento dos parâmetros da matriz de covariâncias em qualquer ponto desejado, inclusive naqueles não observados (Littell *et al.*, 1996). As funções de covariância reduzem o número de parâmetros a serem

FUNÇÕES DE COVARIÂNCIA E CLASSIFICAÇÃO DE REPRODUTORES

estimados, diminuindo o esforço computacional e facilitando a compreensão dos resultados encontrados por parte do pesquisador (Littell *et al.*, 1996; Longford, 1993).

As funções de covariância têm encontrado emprego crescente nas mais diversas áreas (Clark, 1979; Diggle *et al.*, 1994; Kirkpatrick *et al.*, 1990; Lindsey, 1993; Littell *et al.*, 1998; Marx e Thompson, 1987; Meyer, 1998; Meyer, 1999; Meyer, 2000; Meyer e Hill, 1997; Patel, 1986; Ripley, 1981), devendo-se isto principalmente à eficiente implementação da metodologia em pacotes estatísticos de análise de dados, tais como o S-PLUS (MATHSOFT, 1999) e o SAS (SAS, 2000).

Considerando o exposto anteriormente, este trabalho teve por objetivo demonstrar a identificação de uma função de covariância adequada para descrever efeitos dentro de unidades experimentais, utilizando dados de crescimento do tipo peso-idade, de caprinos Moxotó criados no estado de Pernambuco, verificando as consequências da escolha de diferentes funções de covariância sobre a classificação dos reprodutores.

MATERIAL E MÉTODOS

Foram utilizados dados de 412 caprinos da raça Moxotó, filhos de 151 mães e 50 pais, nascidos entre 1984 e 1988, nos meses de março a julho, no município de Serra Talhada, na região do Sertão do estado de Pernambuco. As condições climáticas médias para os meses de nascimento foram: umidade relativa 53,97% ($\pm 11,45$), temperatura 26,18°C ($\pm 1,59$); pluviosidade 89,33mm³ ($\pm 109,29$). As pesagens ocorreram ao nascimento e, posteriormente, aos 28, 112, 196 e 364 dias de idade, totalizando 1749 observações, pois nem todos os animais possuíam todas as observações.

Utilizou-se o seguinte modelo misto, denominado de *modelo completo*, para a descrição dos dados de pesagem:

$$Y_{ijklmnopqrs} = \mu + S_k + A_l + M_m + T_n + P_o + O_p + U + C + H + D + D^2 + R_q(A_l) + F_r + e_{ijklmnopqrs}$$

onde:

$Y_{ijklmnopqrs}$: representa o peso observado para o animal i na idade j , $i = 1, \dots, 412$; $j = 0, 28, 112, 196, 364$;

μ : representa uma constante comum a todas as observações;

S_k : representa o sexo, $k = 1$ (macho), 2 (fêmea);

A_l : representa o ano de nascimento do animal, $l = 1984, 1985, 1986, 1987, 1988$;

M_m : representa o mês de nascimento do animal, $m = 3, 4, 5, 6, 7$;

T_n : representa o tipo de parto, $n = 1$ (simples), 2 (duplo);

P_o : representa o tipo de pelagem, $o = 1$ (simples), 2 (composta);

O_p : representa a ordem de parto, $p = 1, 2, 3, 4, 5, 6$;

U : representa o efeito linear da umidade relativa do ar no dia da pesagem;

C : representa o efeito linear da temperatura no dia da pesagem;

H : representa o efeito linear da pluviosidade no dia da pesagem;

D : representa o efeito linear do fator tempo (dias);

D^2 : representa o efeito quadrático do fator tempo (dias);

$R_q(A_l)$: representa o efeito aleatório do reprodutor q ($q = 1, \dots, 50$) dentro de ano l ;

F_r : representa o efeito aleatório da mãe r ($r = 1, \dots, 151$) e

$e_{ijklmnopqrs}$: representa o erro aleatório associado a cada observação.

Encontrada a matriz de covariância mais adequada (descrito a seguir), as variáveis do modelo passaram por uma seleção do tipo *backward* (Neter *et al.*, 1989), sendo o mesmo processo realizado com a matriz do tipo simetria composta. Este passo teve por objetivo retirar todas os efeitos fixos não significativos do modelo, produzindo o que foi denominado de *melhor modelo* para ambas matrizes de covariância. Este procedimento segue sugestão de Khattree e Naik (1999), Littell *et al.* (1996) e Matsushita (1994), pois na comparação de diferentes estruturas de covariância a eliminação de efeitos no modelo deve ser feita entre os

efeitos fixos e não entre os aleatórios, para não comprometer a estimação dos parâmetros da matriz de covariâncias, pois ao retirar ou incluir efeitos aleatórios, já não estariam sendo comparadas as mesmas matrizes devido a mudança nas suas dimensões, influenciando todos os demais passos da análise.

Não foi composto grupo contemporâneo por estação de nascimento, pois os animais apresentaram nascimento em uma única estação do ano, a época chuvosa para a região, que vai de março a setembro. Optou-se, então, por utilizar os efeitos de ano e mês de nascimento. O mesmo processo foi repetido com a matriz do tipo simetria composta. Os efeitos climáticos são dinâmicos, mudando de valor e influenciando diferentemente os pesos a cada medida, desta forma, seus valores foram incluídos a cada pesagem, tal como sugerido por Verbyla (1988). Utilizou-se o efeito de pai dentro de ano porque se verificou o uso diferenciado de alguns pais em determinados anos.

Para a matriz de covariâncias dos efeitos entre animais foi ajustada uma matriz com configuração de componentes de variância (VC), pois os reprodutores não eram relacionados. Para os efeitos dentro das pesagens foram ajustadas as matrizes com estrutura simetria composta (CS), simetria composta heterogênea (CSH), autorregressiva de ordem um (AR(1)), autorregressiva heterogênea de ordem um (ARH(1)), autorregressiva de médias móveis de ordem um (ARMA(1,1)), Toeplitz heterogênea (TOEPH), não estruturada com correlações (UNR), componente de variância (VC) e *spatial power* (SP(POW)).

O PROC GLM do SAS (SAS, 2000) foi usado para se realizar o teste de esfericidade de Mauchly. Os parâmetros do modelo misto foram estimados por máxima verossimilhança restrita, através do PROC MIXED do programa SAS (SAS, 2000). A possibilidade de diferença significativa entre matrizes que são uma sub-tipo da outra foi realizada através do teste de razão de

verossimilhança, descrita por Khattree e Naik (1999) e Matsushita (1994). A matriz de melhor ajuste foi escolhida pelo melhor valor de ajuste do critério de informação de Akaike (AIC).

Os valores preditos para os efeitos aleatórios obtidos através do melhor modelo com a matriz mais adequada e também com matriz tipo CS foram comparadas em tabela, para demonstrar a influência da escolha da matriz sobre a classificação dos reprodutores (pais e mães) e seus valores preditos.

RESULTADOS E DISCUSSÃO

TESTE DE ESFERICIDADE

Para a amostra estudada o critério de Mauchly para o teste de esfericidade foi de 0,3119374 (g.l.=9, $\chi^2=26,11435$, $p<0,01$), o que indicou a não esfericidade da matriz, devendo-se testar outras configurações para o relacionamento intra-específico.

MATRIZES DE COVARIÂNCIA

A estrutura CSH apresentou bom ajuste, sendo que ela supõe heterogeneidade de variância e também um comportamento não constante das covariâncias entre todas as observações, mas não apresenta um fator de correlação explícito em sua definição. A forma ARH(1), além das mesmas suposições para as covariâncias, possui um fator heterogêneo de correlação, e não de covariância como a CSH. Este fator, mais adequado na descrição de medidas de peso, proporcionou um ajuste ligeiramente superior à estrutura ARH(1) em relação à CSH (**tabela I**).

A matriz do tipo ARH(1) foi considerada a mais adequada, apesar da matriz com o segundo melhor ajuste, a CSH, também apresentar um bom ajuste (**tabela I**). Estruturas que não possibilitam a correlação entre as diferentes medidas, quando elas estão presentes, também apresentaram ajuste ruim, como no caso da estrutura VC. As matrizes que permitem correlação entre as diferentes medidas, mas supõem que elas

FUNÇÕES DE COVARIÂNCIA E CLASSIFICAÇÃO DE REPRODUTORES

sejam constantes não estão em conformidade com o comportamento encontrado na amostra, em que medidas mais próximas apresentam correlações mais altas, assim, as estruturas AR(1) e ARMA(1,1) não apresentaram ajuste superior. O uso de matrizes não estruturadas como a UNR, elevou o número de parâmetros a serem estimados, além de conduzir a problemas de convergência e estimação de parâmetros, apresentando um padrão de comportamento errático para a característica sob estudo, concordando com o discutido por Littell *et al.* (1998).

Marx e Thompson (1987) indicaram que as relações de covariância seriam descritas mais apropriadamente por matrizes dos tipos esférico, exponencial, gaussiano, linear, linear generalizado e logarítmico no caso de dados espacialmente distribuídos. Para o caso onde as repetidas observações são tomadas no tempo e suspeita-se de correlação entre as mesmas, o que é comum em experimentos farmacológicos e de crescimento, um modelo padrão seria uma

matriz autorregressiva de ordem um (AR(1)), empregada para dados com ou sem igual espaçamento entre observações (Davidian e Giltinan, 1995). O software SAS (SAS, 2000) possui ainda outras formas para a matriz de covariâncias, como exemplificado por Khattree e Naik (1999) e Littell *et al.* (1996).

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 serem 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 ser utilizado para comparar modelos com os mesmos efeitos

Tabela I. Resultados dos ajustes dos modelos completos segundo a escolha da matriz de covariâncias para efeito dentro de animais, testes de razão de máxima verossimilhança restrita e critério de informação de Akaike (AIC). (Complete model adjust results by covariance matrix choice to intra-specific effects, restricted maximum likelihood test, and Akaike information criteria (AIC)).

TMC	RLMVR	NPV ^a	GLLRT	TRV ^b	CIA
CS	6639,0	4	3	A	-3323,5
CSH	5373,8	8	8	B	-2694,9
AR(1)	6532,6	4	1	C	-3268,3
ARH(1)	5362,0	8	8	D	-2689,0
ARMA(1,1)	6523,8	5	3	E	-3265,9
TOEPH	8070,0	11	7	F	-4042,0
UNR	7454,6	17	13	G	-3740,3
VC	6647,0	3	3	H	-3326,5
SP-POW	6403,6	4	1	I	-3203,8

TMC: Tipo da matriz de covariância; RLMVR: -2* res. log máxima verossimilhança restrita; NPV: Número de parâmetros de covariância a serem estimados; GLLRT: Graus de liberdade para o teste LRT; TRV: Teste de razão de verossimilhança (LRT); CIA: Critério de informação de Akaike (AIC).

^aRepresenta o número de parâmetros da matriz de covariância mais os fatores aleatórios do modelo.

^bLetras diferentes na mesma coluna indicam matrizes diferentes pelo teste LRT.

fixos, mas diferentes estruturas de variância, sendo que o modelo com o maior 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).

INFLUÊNCIA DE EFEITOS FIXOS

Houve pouca variação climática entre os meses de nascimento e os animais estavam bem aclimatados à região, justificando que os efeitos fixos das variáveis climáticas umidade, temperatura e pluviosidade não foram significativos para nenhuma das estruturas ARH(1) ou CS (**tabela II**). O sexo, tipo de parto, efeitos linear e quadrático de dias foram significativos ($p < 0,01$) para todos os modelos. O efeito de ano de nascimento não foi significativo somente para ARH(1). O tipo de pelagem não foi significativo para o modelo ARH(1).

Em uma maneira geral, os modelos ARH(1) e CS diferiram quanto à significância do efeito fixo de ano de nascimento e tipo de pelagem (**tabela II**).

Observou-se que conforme se alterou a configuração da matriz de covariância,

alteraram-se também as significâncias dos efeitos fixos e aleatórios estudados no modelo, concordando com o exposto por Wolfinger e Chang (1995), que salientaram a importância da especificação da matriz, pois os testes estatísticos seriam uma função dela.

Ao se tratar a curva de crescimento como uma trajetória polinomial, dados pelos efeitos linear e quadrático do tempo em dias, flexibilizam-se as oscilações existentes entre diferentes indivíduos e no próprio indivíduo (Xavier, 2000), melhorando a explicação do comportamento das observações (Wolfinger e Chang, 1995) e ajustando um modelo que reduz o número parâmetros a serem estimados (Von Ende, 1993; Kshirsagar e Smith, 1995).

MELHOR MODELO

O melhor modelo para a descrição dos dados de pesagem, para as matrizes com estrutura ARH(1) e CS, foi:

$$Y_{ijknlrs} = \mu + S_k + T_n + D + D^2 + R_q(A_i) + F_r + e_{ijknlrs}$$

cujos termos já foram descritos (modelo completo), sendo que os valores estimados e preditos para os efeitos fixos e aleatórios

Tabela II. Teste para os efeitos fixos do modelo, segundo a escolha do tipo da matriz de covariâncias, para os modelos completos. (Model fixed effects test, on choice of covariance matrix type, for complete models).

Tipo da matriz	Parâmetro											
	Estatística	S	A	M	T	P	O	D	U	C	H	D2
CS												
GL num.		1	4	4	1	1	5	1	1	1	1	1
GL den.		1258	61	1258	1258	1258	1258	1258	1258	1258	1258	1258
F		18,09***	6,12***	0,64	13,95***	12,7***	0,51	1568,26***	0,01	0,1	0,05	380,27***
ARH(1)												
GL num.		1	4	4	1	1	5	1	1	1	1	1
GL den.		1258	61	1258	1258	1258	1258	1258	1258	1258	1258	1258
F		15,58***	1,32	1,18	18,29***	1,4	0,53	1747,18***	0,86	0,16	0,05	330,39***

*** $p < 0,01$, para o teste F.

FUNÇÕES DE COVARIÂNCIA E CLASSIFICAÇÃO DE REPRODUTORES

Tabela III. Valores preditos (efeitos aleatórios) e estimados (efeitos fixos), para o melhor modelo para as matrizes do tipo ARH(1) e CS. (Random effects predicted values and fixed estimated values, for best model for ARH(1) and CS matrix).

Tipo da matriz	Parâmetro	Valor predito	Erro padrão	Valor de Z	Prob. > Z	AIC	
ARH(1)	R(A)	0,01826	0,01006	1,81	0,0348	-2692,2	
	F	0,04264	0,01332	3,20	0,0007		
	Peso nasc.	0,1542	0,01446	10,67	<0,0001		
	Peso 28 dias	2,1110	0,1672	12,62	<0,0001		
	Peso 118 dias	7,0044	0,5751	12,18	<0,0001		
	Peso 190 dias	11,3904	0,9751	11,68	<0,0001		
	Peso 364 dias	19,1993	1,9827	9,68	<0,0001		
	ARH(1)	0,4252	0,03076	13,82	<0,0001		
CS	R(A)	1,0544	0,2532	4,16	<0,0001	-3376,4	
	F	0,2368	0,1204	1,97	0,0246		
	Fator CS	0,2549	0,1517	1,68	0,0929		
	Resíduo	4,7350	0,1974	23,98	<0,0001		
		Nível	Valor estimado	Erro padrão	GL	Valor de t	Prob. > t
ARH(1)	Intercepto		1,8691	0,04968	65	37,65	<0,0001
	S	1	0,1973	0,04872	1271	4,05	<0,0001
	S	2	0,0000	-	-	-	-
	T	1	0,1680	0,04335	1271	3,88	0,0001
	T	2	0,0000	-	-	-	-
	D	-	0,06614	0,001574	1271	42,02	<0,0001
	D2	-	-0,00009	0,000005	1271	-18,22	<0,0001
	CS	Intercepto		1,7606	0,2088	65	8,43
S		1	0,5836	0,1587	1271	3,68	0,0002
S		2	0,000	-	-	-	-
T		1	0,5622	0,1434	1271	3,92	<0,0001
T		2	0,0000	-	-	-	-
D		-	0,06189	0,001559	1271	39,70	<0,0001
D2		-	-0,00009	0,000004	1271	-19,61	<0,0001

estão demonstrados na **tabela III**.

O valor estimado para o efeito linear do tempo foi próximo para ambos os modelos, e o efeito quadrático foi o mesmo (**tabela III**). Entretanto, para os efeitos sexo e tipo de parto, o modelo ARH(1) apresentou valores inferiores ao do modelo CS. O intercepto para o modelo CS foi inferior ao do modelo ARH(1).

VALORES PREDITOS

Segundo Littell *et al.* (1998) e Wolfinger e Chang (1995), as diferentes estruturas das matrizes de covariância levariam a diferen-

tes valores preditos para os efeitos aleatórios do modelo analisado. A matriz do tipo CS seria menos eficiente que a ARH(1), por desconsiderar a heterogeneidade de variância nas diferentes idades e as diferentes covariâncias entre elas, o que seria denotado pelo menor ajuste (**tabela I**).

Os valores preditos para os reprodutores (**tabela IV**) pela matriz CS apresentaram-se abaixo daqueles da matriz ARH(1), para os anos de 1986 e 1987, acima, para os anos de 1984 e 1988, e próximos, para o ano de 1985. Para as fêmeas o comportamento não teve um padrão definido (**tabela V**).

Tabela IV. Valores preditos para machos por ano, segundo a matriz de covariância ARH(1) ou CS, para os melhores modelos (dez primeiros). (Male predicted values by year, on choice of covariance matrix ARH(1) or CS, for best models (first ten)).

Class	1984			1985			1986		
	C	ARH(1)	CS	C	ARH(1)	CS	C	ARH(1)	CS
1	187	0,1219	2,1221	185	0,1978	800	188	0,08792	800
2	805	0,06283	1,396	808	0,08037	810	804	0,04637	813
3	80	0,05453	1,259	813	0,04335	807	813	0,03539	182
4	802	0,04216	1,1648	809	0,03836	801	814	0,01435	197
5	801	0,02456	1,159	803	0,03098	808	187	0,01283	188
6	188	0,02182	1,1283	807	-0,0001	184	197	-0,00112	184
7	814	0,00125	1,1209	801	-0,00048	809	186	-0,02779	814
8	803	-0,00069	1,025	800	-0,01038	813	800	-0,03985	803
9	806	-0,01203	0,948	814	-0,04063	185	803	-0,05422	801
10	185	-0,03335	0,7563	810	-0,04674	814	811	-0,05708	183

Class	1987			1988		
	C	ARH(1)	CS	C	ARH(1)	CS
1	192	0,2453	0,2106	188	0,04101	207
2	183	0,1702	0,1437	186	0,02448	197
3	194	0,1125	-0,02069	193	0,01002	193
4	184	0,1117	-0,2169	182	-0,01977	186
5	187	0,09569	-0,4112	212	-0,02695	188
6	801	0,08455	-0,5541	191	-0,02831	182
7	804	0,04468	-0,943	207	-0,04534	212
8	191	0,03253	-0,9961	195	-0,05111	199
9	186	0,02633	-1,0125	197	-0,09537	205
10	182	0,0162	-1,0895	194	-0,1014	183

Class: Classificação; C: Código.

FUNÇÕES DE COVARIÂNCIA E CLASSIFICAÇÃO DE REPRODUTORES

Tabela V. Valores preditos para fêmeas, segundo a matriz de covariância ARH(1) ou CS, para os melhores modelos (primeiras quarenta). (Female predicted values, on choice of covariance matrix ARH(1) or CS, for best models (first forty)).

Class	C	ARH(1)	C	CS	Class	C	ARH(1)	C	CS
1	42	0,3253	58	0,6886	21	751	0,1441	21	0,268
2	32	0,2954	91	0,5765	22	109	0,1433	77	0,2594
3	47	0,274	95	0,4845	23	48	0,1295	22	0,2432
4	91	0,2659	86	0,4459	24	45	0,1257	49	0,2394
5	31	0,2485	702	0,409	25	34	0,1187	83	0,2169
6	12	0,2361	31	0,4009	26	101	0,1112	13	0,2023
7	69	0,2291	12	0,3769	27	757	0,1063	62	0,2007
8	94	0,2188	45	0,3754	28	9	0,1006	47	0,1971
9	702	0,2167	30	0,3638	29	95	0,1006	121	0,1926
10	86	0,2078	42	0,3551	30	71	0,09788	110	0,1924
11	715	0,1931	8	0,354	31	61	0,09607	179	0,1772
12	46	0,1916	76	0,3503	32	123	0,09233	106	0,1752
13	58	0,1844	44	0,331	33	100	0,09098	29	0,1739
14	24	0,1825	48	0,319	34	78	0,09037	101	0,1676
15	52	0,1818	109	0,3049	35	15	0,09022	68	0,1651
16	68	0,1743	59	0,3048	36	62	0,08984	39	0,1605
17	772	0,1627	244	0,3031	37	120	0,08745	751	0,153
18	54	0,1613	120	0,2947	38	98	0,07719	73	0,1367
19	73	0,1557	84	0,2715	39	29	0,07271	69	0,12
20	43	0,15	782	0,2713	40	704	0,06666	123	0,1154

Class: Classificação; C: Código.

A mudança no valor predito para os reprodutores dentro dos anos segundo o tipo de matriz (ARH(1) ou CS), tanto pais quanto mães (**tabelas IV e V**), deixou evidente a mudança na classificação dos animais conforme a matriz de covariância utilizada.

Desta forma, a pressuposição, pelo pesquisador, de que a matriz de covariâncias possui estrutura CS, quando outro tipo seria mais plausível, poderia levar a grandes erros de estimação e predição de valores futuros. Devendo-se considerar, ainda, o fato de que muitos pesquisadores ignoram que a matriz de covariâncias assumida por determinados pacotes estatísticos, caso nenhuma forma seja especificada, é a estrutura VC, que é ainda mais simples que a CS, por considerar as covariâncias como nulas e todas variâncias tendo o mesmo valor.

As implicações para programas de melhoramento genético animal são imediatas, uma vez que os valores preditos são a base para a seleção dos animais. Uma vez selecionado o animal errado, poderia ocorrer o comprometimento de todo o programa de melhoramento genético.

Apesar da implementação da metodologia ser relativamente recente nos pacotes estatísticos (SAS, 2000, MATHSOFT, 1999), vários autores recomendam uma análise criteriosa das estruturas de covariância possíveis de serem utilizadas ao invés da aplicação direta de uma estrutura qualquer, sem critérios estatístico-experimentais que a justifiquem (Littell *et al.*, 1998; Longford, 1993; Keselman, 1998). Deve-se ter em mente que a estrutura real da matriz de covariâncias é desconhecida, devendo-se buscar a mais adequada entre as disponíveis para uso através de critérios bem definidos.

CONCLUSÕES

O correto acondicionamento das relações entre efeitos entre e dentro das observações de pesagens, através de uma matriz que admitiu diferentes correlações e covariâncias entre as diferentes pesagens, apresentou resultados mais adequados e satisfatórios que aqueles de uma análise efetuada em um modelo de repetibilidade, comumente utilizado.

A suposição de que a esfericidade está presente na amostra, sem que se efetue o teste adequado, poderia levar o pesquisador a incorrer em erros de significância para os

testes estatísticos, o que causaria conclusões erradas a respeito de quais efeitos seriam realmente influentes ou não, levando também a predição de valores futuros tendenciosos.

Recomenda-se a verificação da adequação da matriz de covariâncias, através do uso da metodologia de modelos mistos utilizando-se estimativas REML, apesar do maior tempo de análise necessário pelo pesquisador, por tratar-se de processo computacionalmente intensivo, pois os resultados seriam mais confiáveis, e as inferências a partir dos mesmos mais seguras.

BIBLIOGRAFIA

- Clark, I. 1979. Practical geostatistics. Applied Science Publications. Essex.
- Davidian, M. and Giltinan, D.M. 1995. Nonlinear models for repeated measurement data. Chapman & Hall. London.
- Diggle, P.G., Liang, K.Y. and Zeger, S.L. 1994. Analysis of longitudinal data. Clarendon Press. Oxford.
- 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 Feldt, L.S. 1970. Conditions under which mean square ratios in repeated measurements designs have exact F distributions. *J. Amer. Stat. Ass.*, 65: 1582-1589.
- Jaffrézic, F. and Pletcher, S.D. 2000. Statistical models for estimating the genetic basis of repeated measures and other function-valued traits. *Genetics*, 156: 913-922.
- Keselman, H.J., Algina, J., Kowalchuk, R.K. and Wolfinger, R.D. 1998. A comparison of two approaches for selection covariance structures in the analysis of repeated measurements. *Comm. Stat., Series B*, 27: 591-604.
- Khattree, R. and Naik, D.N. 1999. Applied multivariate statistics with SAS software. 2nd ed. SAS Institute Inc. Cary.
- Kirkpatrick, M., Lofsvold, D. and Bulmer, M. 1990. Analysis of inheritance, selection and evolution of growth trajectories. *Genetics*, 124: 979-993.
- Kirkpatrick, M. and Heckman, N. 1989. A quantitative genetic model for growth, shape, and other infinite-dimensional characters. *J. Math. Biol.*, 27: 429-450.
- Kshirsagar, A.M. and Smith, W.B. 1995. Growth curves. Marcel Dekker Inc. New York.
- Lindsey, J.K. 1993. Models for repeated measures. Clarendon Press. Oxford.
- Littell, R.C., Freund, R.J. and Spector, P.C. 1991. SAS System for linear models. 3^a ed. SAS Institute Inc. Cary (NC).
- Littell, R.C., Milliken, G.A., Stroup, W.W. and Wolfinger, R.D. 1996. SAS System for mixed models. SAS Institute Inc. Cary (NC).
- Littell, R.C., Henry, P.R. and Ammerman, C.B. 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.
- Malheiros, E.B. 2001. Precisão da análise de experimentos com medidas repetidas usando procedimentos do SAS. *Rev. Mat. Estat.*, 19: 253-272.
- Marx, D. and Thompson, K. 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

FUNÇÕES DE COVARIÂNCIA E CLASSIFICAÇÃO DE REPRODUTORES

- 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. Stat.*, 11: 204-209.
- Meyer, K. 1998. Estimating covariance functions for longitudinal data using a random regression model. *Gen. Select. Evol.*, 30: 221-240.
- Meyer, K. 1999. Estimates of genetic and phenotypic covariance functions for post-weaning growth and mature weight of beef cows. *J. Anim. Breed. Gen.*, 116: 181-205.
- Meyer, K. 2000. Random regressions to model phenotypic variation in monthly weights of Australian beef cows. *Liv. Prod. Sci.*, 65: 19-38.
- Meyer, K. and Hill, W.G. 1997. Estimation of genetic and phenotypic covariance functions for longitudinal data by restricted maximum likelihood. *Liv. Prod. Sci.*, 47: 185-200.
- Milliken, G.A. and Johnson, D.E. 1992. Analysis of messy data. v.1. Designed experiments. Chapman & Hall. New York.
- Neter, J., Wasserman, W. and Kutner, M.H. 1989. Applied linear regression models. Irwin. Homewood.
- Nunes-Anton, V. and Woodworth, G.G. 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.
- Ripley, B.C. 1981. Spatial statistics. John Wiley & Sons. New York.
- SAS. 2000. SAS/STAT User's guide. Version 8. v.2. SAS Institute Inc. Cary (NC).
- Verbyla, A.P. 1988. Analysis of repeated measures design with changing covariates. *Biometrika*, 75: 172-174.
- Von Ende, C.N. 1993. Repeated measures analysis: growth and other time-dependent measures. In: Scheiner, S.M. and J. Gurevitch (eds.). Design and analysis of ecological experiments. Chapman & Hall. New York. pp. 113-137.
- Wolfinger, R. and Chang, M. 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.