

SELEÇÃO DE MODELOS DE REGRESSÃO ALEATÓRIA PARA AJUSTE DE PESOS DO NASCIMENTO AO MOMENTO DA SELEÇÃO (378 DIAS) DE MACHOS NELORE^{1,3}

AUTORES

JOSLAINE NOELY DOS S. G. CYRILLO^{2,5}, MAURÍCIO MELLO DE ALENCAR^{4,5}, ALEXANDER GEORGE RAZOOK^{2,5}, MARIA EUGÊNA ZERLOTTI MERCADANTE², LEOPOLDO ANDRADE DE FIGUEIREDO²

¹ Instituto de Zootecnia - Centro APTA de Bovinos de Corte - Estação Experimental de Zootecnia de Sertãozinho - CP 63 - Projeto IZ - 6.1980011-00 - <http://www.iz.sp.gov.br/eezooser>.

² Pesquisador - Instituto de Zootecnia - Centro APTA de Bovinos de Corte - Estação Experimental de Zootecnia de Sertãozinho - CP 63 - eezooser@izsp.br.

³ Parte da Tese apresentada pelo primeiro autor ao curso de pós-graduação em Zootecnia UNESP - Jaboticabal - SP

⁴ Pesquisador - Embrapa Pecuária Sudeste - São Carlos - mauricio@cppse.embrapa.com.br

⁵ Bolsista do CNPq.

RESUMO

Foram utilizados pesos (17.942 registros) de 2.582 machos da raça Nelore, nascidos de 1976 a 2001, filhos de 186 touros e 1.035 matrizes, provenientes dos rebanhos selecionados (NeS e NeT) e controle (NeC) do Projeto de Melhoramento Genético das Raças Zebu de Sertãozinho. Os dados foram classificados, em uma escala contínua, em intervalos de sete dias de idade, perfazendo 60 idades, com o mínimo de três e o máximo de sete registros por animal. A grande maioria dos animais (96%) apresenta sete registros. Foram testados 21 modelos de regressão aleatória, compostos por diferentes combinações de ordens (k) dos efeitos aleatórios genético aditivo direto (A), de ambiente permanente do animal (C) e de ambiente permanente materno (Q), além do efeito residual, com estruturas de variâncias homogêneas (R1) e heterogêneas, com quatro (R4) ou oito (R8) classes. Os efeitos fixos foram: grupo de contemporâneos (ano de nascimento, rebanho), mês de nascimento e idade da mãe, além da idade do animal ajustado sob polinômio ortogonal de Legendre de ordem três. Os critérios de informação Bayesiano de Schwarz (BIC) e de Akaike (AIC) acusaram os modelos com kA, kC e kQ iguais a 6,5,3 e 6,6,4 com respectivas variâncias residuais divididas em quatro (R4) ou oito classes (R8) como os mais adequados ao ajuste dos dados de machos Nelore do nascimento ao momento da seleção.

PALAVRAS-CHAVE

bovinos de corte, crescimento, dados longitudinais, estruturas de variâncias

TITLE

SELECTION OF RANDOM REGRESSION MODELS FOR ADJUSTMENT OF WEIGHTS FROM BIRTH TO SELECTION AGE (378 DAYS) OF NELLORE MALES

ABSTRACT

Weights (17,942 records) of 2,582 Nelore males, born from 1976 to 2001, sons of 186 bulls and 1.035 cows, from the selected (NeS and Net) and the control (NeC) herds of the Sertãozinho Zebu Breeds Genetic Improvement Project, were used in this study. The data were classified, in a continuous scale, in intervals of seven days of age, with a total of 60 ages, with the minimum of three and maximum of seven records per animal. The great majority of the animals (96%) had seven records. Twenty one random regression models composed by different combinations of orders (k) of the genetic additive direct (A), permanent environmental due to the animal (C), and permanent maternal environment (Q) effects, besides the residual effect, with homogeneous structures of variances (R1) and heterogeneous structures of variance, with four (R4) or eight (R8) classes, were tested. The fixed effects were: contemporary groups (year of birth, herd), month of birth and age of dam, besides the age of the animal adjusted under Legendre's orthogonal polynomial of order three. The Schwarz Bayesian (BIC) and the Akaike (AIC) information criteria accused the models with kA, kC and kQ of orders 6,5,3 and 6,6,4 with

respective residual variances divided in four (R4) or eight classes (R8) as the more appropriate to adjustment of these Nelore data, from the birth to selection age.

KEYWORDS

beef cattle, growth, longitudinal data, structures of variances

INTRODUÇÃO

Muitas características de interesse em melhoramento genético são expressas e podem ser registradas continuamente sobre longos períodos da vida do animal (KIRKPATRICK & HECKMAN, 1989). Essas características são também chamadas longitudinais ou de dimensão infinita. Recentemente, tem-se observado crescente interesse no uso de funções de covariâncias (FC) e modelos de regressão aleatória (MRA) para análise de dados, como o peso corporal em gado de corte, cuja mudança é gradual e contínua ao longo do tempo (ALBUQUERQUE & MEYER 2001; MEYER, 1998a, 1999, 2002; NOBRE, 2001; SAKAGUTI, 2000).

Entre outras vantagens, a metodologia de MRA/FC permite que os parâmetros de uma característica quantitativa sejam descritos como uma função sob uma escala contínua, ocasionando mudança no panorama descontínuo das estimativas de variâncias e covariâncias (KIRKPATRICK et al., 1990). Entretanto, a aplicação de MRA requer número considerável de parâmetros para descrever as funções de covariâncias associadas, principalmente, de funções de covariâncias relacionadas ao crescimento de gado de corte em idades precoces, o qual está sujeito aos efeitos genéticos e de ambiente permanente materno, entre outros (MEYER, 2002).

O processo de seleção de machos raças Zebu de Sertãozinho caracteriza-se por apresentar frequência contínua de coleta de registros de pesos em um teste de desempenho individual (Prova de Ganho de Peso). Nesse contexto, esse estudo teve por objetivo definir modelos de regressão aleatória de ordens adequadas e parcimoniosas para estimar funções de (co)variâncias genéticas e fenotípicas dos pesos obtidos durante os 13 primeiros meses de vida de animais machos Nelore, usando modelos de regressão aleatória.

MATERIAL E MÉTODOS

Foram utilizados pesos (17.942 registros) de 2.582 machos da raça Nelore, nascidos de 1976 a 2001, filhos de 186 touros e 1.035 matrizes, provenientes dos rebanhos selecionados (NeS e NeT) e controle (NeC) do Projeto de Melhoramento Genético das Raças Zebu de Sertãozinho. Detalhes do projeto estão descritos em MERCADANTE et al., 2003. Os dados foram classificados em intervalos de sete dias de idade, perfazendo 60 idades, com o mínimo de três e máximo de sete registros por animal. A grande maioria dos animais (96%) apresentou sete registros (Figura 1).

Vinte e um modelos de regressão aleatória foram testados e foram compostos por diferentes combinações de ordens (k) dos efeitos aleatórios genético aditivo direto (A), de ambiente permanente do animal (C) e de ambiente permanente materno (Q), além do efeito residual, com estruturas de variâncias homogêneas (R1) e heterogêneas, com quatro (R4) ou oito (R8) classes.

Assumindo-se que a ordem máxima dos polinômios deve ser menor ou igual ao número de registros para a maioria dos animais (MEYER, 1998b) e, conforme relatado por ALBUQUERQUE e MEYER, 2001, em que polinômios de ordens máximas iguais a seis foram suficientes para descrever a variação genética e fenotípica em dados de animais da raça Nelore nas fases iniciais do crescimento, foi definida inicialmente como ordem máxima para qualquer efeito aleatório $k=6$ (kA, kC e kQ) ou ordem quártica. O número de parâmetros a serem estimados para cada efeito aleatório considerado no modelo é dado por $k(k+1)/2$, em que k é a ordem dos coeficientes de regressão aleatória.

As ordens de ajuste dos polinômios para cada efeito foram suprimidas à medida que as estimativas de variâncias dos coeficientes de regressão aleatória apresentavam variâncias de baixa magnitude e autovalores correspondentes inferiores a 1.

As análises foram feitas utilizando-se o módulo DXMRR presente no programa computacional "DFREML" de MEYER, 1998b. Os efeitos fixos foram: grupo de contemporâneos (ano de nascimento e rebanho), mês de nascimento e idade da mãe, além da idade do animal ajustada sob

polinômio ortogonal de Legendre de ordem três.

As comparações entre os vários modelos foram efetuadas em duas etapas: dentro de estruturas de variâncias residuais e, posteriormente, fixando algumas ordens de ajuste para os efeitos aleatórios e fazendo variar as estruturas residuais. Essas comparações foram feitas utilizando-se o teste de razão de verossimilhança (LRT), que permite comparações entre modelos aninhados, e os critérios de informação de Akaike (AIC) e Bayesiano de Schwarz (BIC), que permitem comparações entre modelos não aninhados (LITTEL et al., 1996).

RESULTADOS E DISCUSSÃO

As médias de peso aumentaram linearmente com a idade, variando de 31,4 a 320,8 kg ao nascer e aos 378 dias de idade, respectivamente (Figura 1). Os desvios-padrão apresentaram tendência similar e o coeficiente de variação diminuiu com a idade.

O resumo das análises de regressão aleatória é apresentado na Tabela 1. Para as análises com estruturas homogêneas de variância do resíduo (R1), os modelos menos parametrizados (M = 1 e 2) não se ajustaram aos dados, apresentando maiores variâncias residuais. O critério de informação BIC acusou como melhor o modelo que considerou os polinômios de ordens $k_A=6$, $k_C=5$ e $k_Q=3$ ($k=6,5,3R1$) com 43 parâmetros estimados e o critério AIC, o modelo de ordens $k=6,6,4R1$, com 53 parâmetros.

Os modelos nos quais foram aplicadas estruturas heterogêneas de variâncias dos resíduos (Modelos 10 a 16), divididas em quatro classes distintas (R4), não apresentaram diferenças significativas pelo LRT ($P<0,05$) (Tabela 1), e os critérios de informação BIC e AIC novamente apontaram os modelos $k=6,5,3R4$ e $k=6,6,4R4$, respectivamente, como os de melhor ajuste dos dados.

As diferentes ordens de ajuste dos modelos (Modelos 17 a 21) com oito classes de variâncias residuais (R8), a exemplo dos anteriores, não diferiram significativamente entre si pelo LRT. Pelos critérios BIC e AIC, as ordens dos efeitos aleatórios que melhor ajustaram os dados foram $k=6,5,3R8$ e $k=6,6,4R8$, respectivamente.

Comparações entre modelos de diferentes classificações do resíduo foram efetuadas entre os modelos apontados como os melhores pelos critérios BIC e AIC e que não diferiram significativamente entre si, dentro de classificação de resíduos, pelo LRT. Segundo o critério BIC, considerando todos os modelos e estruturas residuais, o modelo $k=6,5,3R4$ com 46 parâmetros seria suficiente para modelar as mudanças na variação de pesos no período do nascimento ao momento da seleção. Já o critério AIC apontou o modelo $k=6,6,4R8$, com 60 parâmetros a serem estimados, como o melhor modelo.

Observa-se que os modelos definidos como melhores pelos critérios BIC ($k=6,5,3$) e AIC ($k=6,6,4$) foram os mesmos, independente da estrutura de variâncias do resíduo considerada, e que a divisão dos resíduos em oito classes, e conseqüente aumento de parâmetros, foi penalizada pelo BIC, que apontou, na segunda fase das comparações, o modelo $k=6,5,3R4$ (46 parâmetros) como o de melhor ajuste. Os critérios de informação AIC e BIC não estão de acordo sobre qual modelo é o melhor, mas ambos mostraram que medidas heterogêneas de variância residual são mais adequadas para ajustar dados longitudinais. Segundo HUISMAN et al., 2002, quando se considera homogeneidade, assumindo que ambiente temporário afeta todas as idades da mesma forma, a variação residual é retida, em grande parte, na variância de ambiente permanente do animal, dificultando a decomposição das fontes de variação aleatórias. O melhor caminho seria modelar o termo do erro separadamente e independentemente para cada idade. Entretanto, esse procedimento é oneroso em termos computacionais, o que dificultaria a utilização dessa metodologia em grandes conjuntos de dados.

CONCLUSÕES

Os modelos de regressão aleatória de melhor ajuste para pesos de bovinos Nelore nas fases iniciais

do crescimento incluíram ordens (k) iguais a 6, 5 e 3 e 6, 6 e 4 para os efeitos genético direto, de ambiente permanente de animal, e de ambiente permanente materno, com o resíduo dividido em quatro ou oito classes de variâncias, respectivamente.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ALBUQUERQUE, L.G.; MEYER, K. . Estimates of covariance functions for growth from birth to 630 days of age in Nelore cattle. *J. Anim. Sci.*, v.79, p. 2776-2789, 2001.
2. HUISMAN, A.E.; VEERKAMP, R.F.; ARENDONK, J.A.M. Genetics parameters for various random regression models to describe the weight data of pigs. *J. Anim. Sci.* v. 80, p. 575-582, 2002.
3. KIRKPATRICK, M.; HECKMAN, N. . A quantitative genetic model for growth, shape and other infinite-dimensional characters. *J. Math. Biol.*, v.27, p. 429-450, 1989.
4. KIRKPATRICK, M.; LOFSVOLD, D.; BULMER, M. . Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics*, v. 124, p. 979-993, 1990.
5. LITTEL, R.C.; MILLIKEN, G.A.; STROUP, W.W.; WOLINGER, R.D. SAS system for mixed models. Cary: SAS Institute, 433p., 1996.
6. MERCADANTE, M.E.Z.; PACKER, I.U.; RAZOOK, A.G.; CYRILLO, J.N.S.G., FIGUEIREDO, L.A. . Direct responses to selection for yearling weight on reproductive performance of Nelore cows. *J. Anim. Sci.*, v. 81, p.376-384, 2003.
7. MEYER, K. . Estimating covariance functions for longitudinal data using a random regression model. *Genetic. Sel. Evol.*, v.30, p.221-240, 1998a.
8. MEYER, K. . "DXMRR" - A program to estimate covariance functions for longitudinal data by REML. In: 6th WORLD CONGRESS OF GENETICS APPLIED TO LIVESTOCK PRODUCTION, 6., 1998, Armidale. Proceedings... Armidale: University of New England, 1998b. CD-ROM.
9. MEYER, K. . Estimates of genetic and phenotypic covariance functions for postweaning growth and mature weight of beef cows. *J. Anim. Breed. Genet.* v.116, p.181-205, 1999.
10. MEYER, K. . Estimates of covariance functions for growth of Australian Beef Cattle from a large set of field data. In: 7th WORLD CONGRESS OF GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier. Proceedings... Montpellier: [s.n.], 2002. CD-ROM.
11. NOBRE, P.R.. Analyses of sequential weights of Nelore cattle using multiple trait and random regression models. 2001. 137f. Tese (Doutorado) – Universidade Federal de Viçosa, Viçosa, 2001.
12. SAKAGUTI, E.S.. Funções de covariâncias e modelos de regressão aleatória na avaliação genética do crescimento de bovinos jovens da raça Tabapuã. 2000. 86 f. Tese (Doutorado em Zootecnia) - Universidade Federal de Viçosa, Viçosa, 2000.

Tabela 1 - Ordem das funções de covariâncias (k_A , k_C e k_Q), número de parâmetros estimados (Par.), variância residual (σ_e^2) (Resíduo 1), valores dos critérios de informação BIC e AIC (-10.000,00) e valores de Log L para os modelos testados, teste de razão de verossimilhança (LRT) e valor da probabilidade de P ($P < 0,05$) de acordo com a estrutura da variância dos resíduos

Modelo	K_A	k_C	k_Q	Par.	σ_e^2	BIC ^a	AIC ^a	Log L	LRT	P
Resíduo 1										
1	4	4	4	31	59,900	6146,27	5904,62	-52921,31	(2-1)	0,00
2	5	5	5	46	56,940	5850,62	5492,05	-52700,03	(3-2)	0,00
3	6	5	3	43	51,986	5336,53	5001,34	-52457,67	(4-3)	0,26
4	6	5	4	47	51,936	5365,17	4998,81	-52452,41	(5-4)	0,99
5	6	5	5	52	52,032	5413,25	5008,00	-52451,96	(8-5)	0,02
6	6	6	3	49	52,009	5371,25	4989,31	-52445,65	(7-6)	0,20
7	6	6	4	53	51,983	5398,86	4985,43	-52439,72	(8-7)	0,13
8	6	6	5	58	52,048	5442,58	4990,48	-52437,24	(9-8)	0,99
9	6	6	6	64	52,048	5501,24	5002,37	-52437,18		
Resíduo 4										
10	6	5	5	55		4638,57	4209,85	-52049,93	(15-10)	0,35
11	6	5	3	46		4564,61	4206,05	-52057,03	(12-11)	0,44
12	6	6	3	52		4611,75	4206,42	-52051,21	(14-12)	0,29
13	6	5	4	50		4594,57	4204,82	-52052,41	(14-13)	0,42
14	6	6	4	56		4640,61	4204,10	-52046,05	(15-14)	0,27
15	6	6	5	61		4683,94	4208,45	-52043,22	(16-15)	0,83
16	6	6	6	67		4742,95	4220,70	-52043,09	(16-15)	
Resíduo 8										
17	6	6	3	56		4618,89	4182,38	-52035,19	(20-17)	0,24
18	6	5	3	50		4574,56	4184,82	-52042,41	(19-18)	0,28
19	6	5	4	54		4603,68	4182,75	-52037,38	(20-19)	0,26
20	6	6	4	60		4647,14	4179,45	-52029,72	(21-20)	0,99
21	6	6	6	71		4749,12	4195,68	-52026,84		

^a Valores em negrito correspondem ao modelo escolhido pelo respectivo critério de informação.

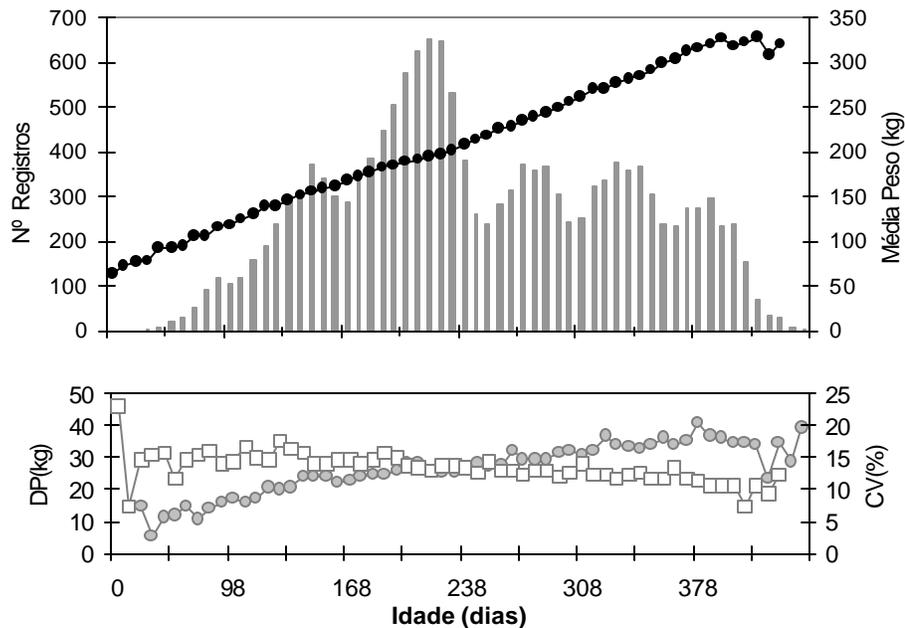


Figura 1 – Distribuição do número de registros (barras), médias de pesos (●), desvios-padrão (○) e coeficientes de variação(□), de acordo com a idade.