



43ª Reunião Anual da Sociedade Brasileira de Zootecnia
24 a 27 de Julho de 2006
João Pessoa - PB

COMPARAÇÃO ENTRE MODELOS DE REGRESSÃO ALEATÓRIA UTILIZANDO FUNÇÕES POLINOMIAIS DE LEGENDRE PARA O AJUSTE DA PRODUÇÃO DE LEITE NO DIA DO CONTROLE DE PRIMEIRAS LACTAÇÕES DE VACAS GIROLANDO (1)

LUCIARA CELI DA S. CHAVES(2), CLAUDIO VIEIRA DE ARAUJO(3), ARY FERREIRA DE
FREITAS(4), ROBLEDO DE A. TORRES(5), PAULO SÁVIO LOPES(5), ALEX SANDRO
SCHIERHOLT(2)

(1)Fapemig-Fundo de Amparo à Pesquisa do Estado de Minas Gerais

(2)Estudante de Doutorado da Universidade Federal de Viçosa-UFV

luciaracelichaves@yahoo.com.br

(3)Professor Adjunto da Universidade Federal Rural da Amazônia-UFRA

(4)Pesquisador da Embrapa Gado de Leite, Juiz de fora, MG

(5)Professor Adjunto da Universidade Federal de Viçosa-UFV

RESUMO

Registros de controles individuais de produção de leite de vacas Girolando foram utilizados na comparação entre modelos de regressão aleatória. Os modelos diferiram entre si por meio da função utilizada para descrever a trajetória da curva de lactação, sendo estas funções polinomiais de Legendre com três, quatro e cinco parâmetros, empregando diferentes números de classes residuais para o efeito de ambiente temporário. Comparações entre os modelos foram realizadas por meio do Teste da Razão de Verossimilhança e AIC. Verificou-se que a medida que se aumentou o número de parâmetros, menor foi o número de classes residuais necessárias para melhor ajuste a descrição dos dados. Para os modelos polinomiais com três parâmetros, o que promoveu o melhor ajuste empregou seis classes residuais, para o modelo com quatro parâmetros, quatro classes residuais e, o modelo que empregou cinco parâmetros, somente duas classes residuais foram suficientes. O modelo polinomial com cinco parâmetros foi o que mostrou menor diferença entre as variâncias de efeito de ambiente temporário localizadas nos extremos da trajetória de lactação, mostrando que a maior heterogeneidade dessas variâncias pode decorrer do uso de funções inadequadas para descrever a trajetória. Pelos resultados do AIC, o modelo polinomial com cinco parâmetros promoveu melhor ajuste para descrição da variação da produção no dia de controle quando comparado aos demais modelos.

PALAVRAS-CHAVE

Curva de lactação, Avaliação Genética

COMPARISON AMONG RANDOM REGRESSION MODELS USING LEGENDRE POLYNOMIALS FUNCTIONS FOR THE ADJUSTMENT OF MILK YIELD IN TEST DAY CONTROLS OF FIRST LACTATION OF GIROLANDO COWS

ABSTRACT

Records of test day milk yield of Girolando cows were used in the comparison of Random Regression Models. The models differed to each other by means of the function used to describe the trajectory of

the lactation curve, being these Legendre polynomial functions with three, four and five parameters, using different numbers of residual classes to temporary environmental effect. The comparisons among the models were based on the Likelihood Ratio Test and the AIC test. It was verified that, as the number of parameters increased, minor were the numbers of residual classes to best adjust the data description. For the polynomial models with three parameters, the one that best adjust the data used six residual classes, for the model with four parameters, four residual classes and, the model that used five parameters, only two residual classes were enough. The polynomial model using five parameters was also the one that showed smaller difference in the variance of temporary environmental effect on the extremes of the lactation curve, showing that the higher heterogeneity of these variances may occur when using function that are inadequate to describe the trajectory. Based on the AIC results, the polynomial model with five parameters promoted better adjustment to the description of the variance in the test day yield when compared with the others models.

KEYWORDS

Lactation curve, Genetic evaluation

INTRODUÇÃO

Os Modelos de Regressão Aleatória (MRA) vêm sendo propostos para a avaliação genética de característica longitudinais como produção de leite no dia do controle (“test day”), e curvas de crescimento. Os MRA têm sido propostos como uma alternativa para modelar tais características (Schaeffer e Dekkers, 1994), pois, permitem a predição de valores genéticos como um todo, para qualquer ponto desejado na escala de tempo utilizada, não havendo necessidade de se criar classes de desempenho arbitrárias ou de utilizar fatores de ajuste para determinada idade (Albuquerque, 2004) e permitem obter estimativas de (co)variâncias para as características (idades) na qual não há observações (por meio de interpolação) (Kirkpatrick e Heckman, 1989). Além disso, têm-se uma melhor utilização dos dados, posto que todas as medidas do animal e de seus parentes são utilizadas na avaliação do mesmo, promovendo, assim, um potencial aumento da acurácia de seleção (Albuquerque, 2004). Entretanto, apesar de reduzir o número de parâmetros em relação às análises convencionais, o emprego de MRA envolve, em geral, grande número de parâmetros correlacionados entre si, o que restringe a sua aplicação em grandes arquivos de dados.

Este trabalho tem como objetivo comparar funções polinomiais de Legendre com diferentes números de parâmetros e diferentes classes para o efeito de ambiente temporário, utilizadas para descrever as trajetórias de lactação de animais Girolando.

MATERIAL E MÉTODOS

O arquivo analisado continha 6.604 registros de controles individuais de produção de leite no dia de controle, provenientes de primeiras lactações de 876 vacas de 4 composições raciais (1/2 Holandês:Gir, 5/8 Holandês:Gir, 3/4 Holandês:Gir e 7/8 Holandês:Gir.), distribuídas em 40 rebanhos e 196 classes de RAE (rebanho-ano-estação).

Os modelos empregados para descrever as variações pertinentes aos dados, diferiram entre si pelo número de parâmetros da função polinomial de Legendre, que variou de três a cinco ($k=3, 4$ e 5).

Os números de dias em lactação (DIM) foram padronizados entre -1 e 1 (DIMP), pela expressão apresentada em Kirkpatrick et al. (1990);

Os modelos de regressão aleatória, aplicados aos registros de produção de leite no dia do controle, consideraram classes de rebanho-estação-ano e composição racial como efeitos fixos, idade ao primeiro parto (efeito linear e quadrático) como covariável. Estações de controle foram definidas como águas (de outubro a março) e seca (de abril a setembro). Os coeficientes de regressão para modelar a trajetória média comum a todos os indivíduos e relacionados com os efeitos genético aditivo e de ambiente permanente, foram obtidos por meio polinômios de Legendre, variando de três a cinco

parâmetros. Os componentes de (co)variâncias foram obtidos pelo método da Máxima Verossimilhança Restrita, por meio do DFREML (MEYER, 1998).

Com o objetivo de verificar a presença de heterogeneidade de variâncias para o efeito de ambiente temporário, este foi considerado constante ao longo de toda a trajetória de lactação em cada modelo, e posteriormente, comparado em situações em que este mesmo efeito foi alocado em duas, quatro e seis classes residuais. As classes residuais foram dispostas como:

Duas classes residuais: de 5 a 230 dias e de 231 a 305 dias de lactação;

Quatro classes residuais: de 5 a 80 dias, de 81 a 155 dias, de 156 a 230 dias e de 231 a 305 dias de lactação;

Seis classes residuais: de 5 a 55 dias, de 56 a 105 dias, de 106 a 155 dias, de 156 a 205 dias, de 206 a 255 dias e de 256 a 305 dias de lactação.

Os diferentes modelos de regressão aleatória que utilizaram a mesma função e mesmo número de parâmetros foram comparados entre si pelo Teste da Razão de Verossimilhança e pelo Critério de Informação de Akaike.

A estatística do teste da razão de verossimilhança é dada por $LR_{ij} = 2\text{Log}L_{i} - 2\text{Log}L_{j}$, em que $\text{Log} L$ é o logaritmo do valor da função de verossimilhança. O teste apresenta distribuição Qui-quadrado, com $n-p$ (diferença entre o número de parâmetros) graus de liberdade. O critério de Informação de Akaike é expresso como: $AIC = 2\text{Log} L + 2p$, em que L é o valor da função de verossimilhança e p é o número de parâmetros estimados.

RESULTADOS E DISCUSSÃO

Verifica-se que para os modelos polinomiais de Legendre com três parâmetros, quanto maior o número de classes empregadas para descrever o efeito de ambiente temporário, melhor foi o ajuste do modelo para descrever os dados. Conforme a Tabela 1, o modelo que promoveu o melhor ajuste foi o que empregou seis classes. À medida que se aumentou o número de parâmetros da função polinomial, menor foi o número de classes residuais necessárias para promover o melhor ajuste. El Faro (2002) encontrou resultados que demonstraram que, ao assumir homogeneidade de variâncias residuais, todos os modelos testados mostraram-se inadequados.

Quando a função polinomial utilizou quatro parâmetros, quatro classes de resíduos foram suficientes. Ao passo que, para a função que empregou cinco parâmetros, foram necessárias somente duas classes de resíduo (Tabela 1). Em conformidade com Albuquerque (2003), que também verificou que se aumentando o grau do polinômio diminui-se a variância residual. Tais resultados revelam que funções com menor número de parâmetros, podem ser insuficientes para separar corretamente as diferentes causas de variação da variação total, inflacionando o efeito de ambiente temporário. Portanto, à medida que a função se torna mais eficiente para descrever a variação dos dados, menor é a necessidade de se utilizar várias classes de resíduo.

Pelos resultados do AIC, o modelo polinomial com cinco parâmetros (17.938,9454) promoveu melhor ajuste para descrição da variação da produção no dia de controle quando comparado aos demais modelos.

As variâncias estimadas para o efeito de ambiente temporário pelos modelos que empregaram funções polinomiais de Legendre com três, quatro e cinco parâmetros, considerando variâncias constantes, dispostas em duas, quatro e seis classes residuais, ao longo dos dias em produção, podem ser observadas na Tabela 2. Independentemente do número de parâmetros da função polinomial, de forma geral, quando se aumentou o número de classes residuais, as estimativas no início da trajetória da lactação aumentaram, enquanto no final diminuíram. À medida que o modelo empregou funções com maior número de parâmetros, houve redução nas estimativas de variâncias.

No modelo que empregou a função polinomial com três parâmetros e seis classes residuais, a diferença entre as variâncias estimadas de ambiente temporário nos extremos da trajetória de lactação foi igual a $2,09 \text{ kg}^2$. Para o modelo que utilizou a função polinomial com quatro parâmetros, quatro classes de

resíduo foram suficientes, a diferença entre as variâncias localizadas nos extremos da trajetória foi igual a $0,98 \text{ kg}^2$, enquanto que para o modelo que empregou a função polinomial com cinco parâmetros, foram necessárias somente duas classes de resíduo, a diferença entre variâncias localizadas nos extremos da trajetória foi igual a $0,89 \text{ kg}^2$. Esse resultado revela que, quanto maior o número de parâmetros na função polinomial que descreve a variação dos dados houve melhor discriminação do efeito de ambiente temporário, reduzindo a necessidade de se alocá-lo em maior número de classes de resíduo. Assim, maior heterogeneidade de variâncias do efeito de ambiente temporário, pode ser decorrência do uso de funções inadequadas para descrever a variação da trajetória.

As estimativas de herdabilidade obtidas em cada modelo com diferentes ordens de ajuste do polinômio e diferentes números de classes de efeito temporário são demonstradas na Figura 1. A formação de diferentes classes de resíduos para o efeito de ambiente temporário, não proporcionou grandes alterações na estimativa de herdabilidade nos modelos com diferentes ordens de ajuste. Comparando as estimativas entre os modelos com diferentes ajustes de polinômios, verifica-se que os modelos com quatro e cinco parâmetros, apresentaram comportamentos bem similares.

No início da lactação, o modelo com três parâmetros proporcionou, principalmente, maiores estimativas de variância genética aditiva e menores estimativas de variância para o efeito de ambiente permanente, em relação aos modelos com maior número de parâmetros na função, proporcionando maiores estimativas de herdabilidade, do quinto até aproximadamente aos 35º dia de produção. Apresentando comportamento inverso nos intervalos de 35 aos 155 dias e de 275 aos 305 dias de produção. As estimativas de herdabilidade entre todos os modelos foram próximas no período de 155 a 275 dias, aproximadamente.

CONCLUSÕES

O modelo que empregou a função polinomial de Legendre com cinco parâmetros e duas classes residuais para o efeito de ambiente temporário foi o mais eficiente em descrever esta variação, minimizando a variância residual e aumentando a precisão do ajuste da função de verossimilhança para descrição das variações presentes nos registros de produção de leite no dia de controle de animais Girolando.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALBUQUERQUE, L.G. Modelos de dimensão infinita aplicados a características de crescimento de bovinos da raça Nelore. Jaboticabal: FCAV/UNESP, 2003. 83p. Tese (Livre Docência) – Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, 2003.
- ALBUQUERQUE, L.G. Regressão Aleatória: nova tecnologia pode melhorar a qualidade das avaliações genéticas. In: SIMPÓSIO DA SOCIEDADE BRASILEIRA DE MELHORAMENTO ANIMAL, 5., 2004, Pirassununga, São Paulo. Anais... Pirassununga: SBMA, 2004. (CD-ROM).
- EL FARO, L. Estimação de componentes de (co)variância para produção de leite no dia do controle de primeiras lactações de vacas Caracu, aplicando-se “test-day models” de dimensão finita e modelos de regressão aleatória. Jaboticabal: FCAV/UNESP, 2002. 102p. Tese (Doutorado em Produção Animal) – Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, 2002.
- KIRKPATRICK, M.; HECKMAN, N. A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. J. Math. Biol., v.27 p.429-450, 1989.
- KIRKPATRICK, M.; LOFSVOLD, D.; BULMER, M. Analysis of the inheritance, selection and evolution of growth trajectories. Genetics, v.124, n.2, p. 979-993, 1990.
- MEYER, K. “DXMRR” – A program to estimate covariance functions for longitudinal data by REML. In: PROCEEDINGS 6TH WORLD CONGRESS GENETIC APPLIED LIVESTOCK PRODUCTION. 1998, Armidale, New South Wales, Australia. Anais...Armidale:8WCGALP, 1998. (CD ROM).