

APLICAÇÃO DE MODELO DE REGRESSÃO ALEATÓRIA PARA AVALIAÇÃO GENÉTICA DA PRODUÇÃO DE LEITE NA RAÇA HOLANDESA¹

**Cláudio Vieira de ARAÚJO²
Robledo de Almeida TORRES³
Cláudio Napolis COSTA⁴
Simone Inoe ARAÚJO⁵
Luciara Celi CHAVES⁶
Ana Paula INOE⁷**

RESUMO: Registros de produção de leite de 68 523 controles leiteiros de 8 536 vacas da raça Holandesa, distribuídas em 266 rebanhos, com parições entre 1996 a 2001, foram utilizados na comparação entre modelos de regressão aleatória para predição de valores genéticos. Os rebanhos foram classificados em níveis de produção alto e baixo de acordo com o desvio padrão para produção de leite. Os modelos de regressão aleatória incluíram os efeitos de rebanho-data do controle, composição genética dos animais; frequência de ordenhas. Utilizaram-se os polinômios de Legendre de segundo grau para descrever a trajetória da curva de lactação. Os modelos diferiram entre si pela adição do nível de produção na formação do grupo de contemporâneos. Para a comparação entre os modelos, utilizaram-se as correlações de Pearson e de Spearman entre os valores genéticos preditos em ambas situações, selecionando 10; 20; 30; 40; 50 e 100% dos melhores reprodutores que apresentaram filhas em ambos níveis de produção. Para todas as situações, o valor genético predito médio foi menor quando se considerou a idade-estação ao parto e nível de produção na formação de grupo de contemporâneos, porém as correlações foram todas próximas da unidade, indicando que seria irrelevante considerar este efeito no modelo de avaliação genética.

TERMOS PARA INDEXAÇÃO: Bovinos Leiteiros, Predição de Valores Genéticos, Seleção.

¹ Aprovado para publicação em 01.07.05

Pesquisa financiada com recursos do CNPq.

² Zootecnista, Dr., Professor da Universidade Federal Rural da Amazônia. Instituto da Saúde e Produção Animal – UFRA/ISPA. Belém (PA). E-mail: araujocv@bol.com.br

³ Engenheiro Agrônomo, Dr., Professor da Universidade Federal de Viçosa (MG).

⁴ Zootecnista, Dr., Pesquisador da Embrapa Gado de Leite, Juiz de Fora (MG).

⁵ Engenheira Agrônoma, Professora Adjunta da UFRA

⁶ Engenheira Agrônoma, M.Sc., Professora da UFRA

⁷ Médica Veterinária, Professora Titular da UNIPAR (PR)

RANDOM REGRESSION MODEL FOR GENETIC EVALUATION OF THE MILK YIELD IN HOLSTEIN BREED

ABSTRACT: Data of 68.523 test day milk yield from 8.536 cows of the Holstein breed, distributed in 266 herds, calving from 1996 to 2001, were used to compare random regression models for predicting breeding values of the sires. The statistical model included the effects of herd-year of calving; season-age of calving; frequency of milking; effect of the genetic group, linear effect of milk production (measured in days from the parity to the day of control) and the random genetic additive effect of the animal. The herds were classified in two production levels, high and low. The random regression models used the second order Legendre's polynomials in order to predict breeding values for the milk yield. In the second model, the contemporary group was composed by age-season of parity and the level of production of the herds. It was used the Pearson and Spearman's correlations between the predicted breeding values in both situations and selecting 10, 20, 30, 40, 50 and 100% of the best sires that had daughters in both levels for comparison of the two models. In each situation, the average breeding value was smaller when it was considered the age-season of parity and production level in the contemporary group. However, all the correlations were close to the unit, indicating that it would be irrelevant to consider this effect in the genetic evaluation model.

INDEX TERMS: Breeding Value Predicted, Dairy Cattle, Selection.

1 INTRODUÇÃO

As avaliações genéticas de reprodutores e vacas das raças leiteiras sempre consideraram, nos últimos anos, a produção aos 305 dias de lactação. A razão da utilização das produções aos 305 dias de lactação se deve ao fato de que o período de lactação ideal em raças leiteiras especializadas deve estar em torno dos 10 meses.

Os modelos de "test-day" são definidos como um procedimento estatístico que considera todos os efeitos genéticos e ambientais relacionados com a produção do dia de controle. Nos últimos anos, o interesse pelos modelos de "test-day" vem aumentando por parte dos melhoristas, uma razão para este interesse é a tendência dos países produtores de leite em reduzir os custos com registros de controle leiteiro.

Os modelos de "test-day", além de predizerem a produção total mais acuradamente, não necessitam de estender a lactação por meio de fatores de ajustamento para animais que não completaram a lactação, o que pode levar a vícios nas predições. Outra razão para o uso da produção no dia de controle, refere-se à maior confiabilidade para as predições dos valores genéticos de touros e vacas, uma vez que há aumento no número de observações por animal, proporcionando a seleção mais precoce dos animais, e, conseqüentemente, diminuindo o intervalo de gerações. Um outro método estatístico sob modelo de "test-day" é o modelo de regressão aleatória (RRM). Neste procedimento são obtidos dois conjuntos de regressões das produções no estágio de lactação. O primeiro conjunto se refere ao

ajuste da curva de lactação para vacas pertencentes a mesma subclasse de efeitos fixos, sendo, portanto, uma curva obtida por uma equação de regressão fixa. O outro conjunto de regressão é considerado aleatório, pois considera os desvios da curva de lactação do animal, em relação à curva de lactação fixa para a subclasse da qual pertence o animal.

Henderson Junior (1982) foi quem deu o primeiro passo para o desenvolvimento dos modelos de regressão aleatória, demonstrando um procedimento para estimação e teste de hipóteses para análise de covariância por meio de modelos mistos, apresentando o caso em que coeficientes de regressão aleatória são incluídos na análise. Surgindo então daí o termo regressão aleatória.

A utilização dos RRM, em relação aos modelos de características múltiplas tradicionais e às funções matemáticas, apresentam uma série de vantagens, tais como: não exigem número mínimo de medidas por animal, ao contrário dos modelos que consideram a produção em 305 dias (JAMROZIK; SCHAEFFER; DEKKERS, 1997); consideram o parentesco existente entre os animais, ao contrário das funções que descrevem a curva de lactação; não necessitam de empregar ajustes (arbitrários) em medidas (controles leiteiros) consideradas como padrão (MEYER; HILL, 1997); permitem a utilização de dados de animais com apenas uma observação (SCHAEFFER; DEKKERS, 1994); permitem a estimação de componentes de

(co)variância entre as produções média do controle, incluindo períodos em que estes não foram mensurados; as análises abrangem todo o intervalo contínuo em que as medidas foram tomadas, permitindo predizer o valor genético dos animais e estimar a persistência da lactação, em qualquer ponto do intervalo em que foram feitas as mensurações fenotípicas; permite estimação de um menor número de parâmetros, principalmente, quando se objetiva obter parâmetros genéticos entre várias características, analisadas simultaneamente (VEERKAMP; GODDARD, 1998; POOL; MEUWISSEN, 1999).

Nos diversos modelos e metodologias de avaliação genéticas, quando se admitem variâncias constantes entre os rebanhos, na predição dos valores genéticos dos animais desconsiderando-se o nível de produção ou a variância estimada entre rebanhos, pode haver uma classificação errônea do mérito genético dos animais. Quando as variâncias aumentam com a média da produção, as produções de filhas de reprodutores criadas em rebanhos mais variáveis influenciariam mais a avaliação dos reprodutores, que as produções de filhas criadas em rebanhos menos variáveis. Quando essa heterogeneidade de variância for de origem ambiental, os reprodutores poderiam sendo selecionados em razão do melhor ambiente em que suas progênes são criadas, do que de seus próprios méritos genéticos (COSTA, 1998; TORRES, 1998).

Variâncias heterogêneas entre rebanhos e aumento da variância, de acordo

com o aumento dos níveis de produções dos rebanhos, regiões ou países, têm sido verificados por vários autores, com Hill et al. (1983), Meyer (1987), Boldman e Freeman (1990), Dong e Mao (1990), Stanton et al. (1991), Dodenhoff e Swalve (1998), Costa (1998), Torres (1998) e Engel et al. (1999).

Quando a heterogeneidade de variância é desprezada, embora esteja presente, a produção das filhas de determinado reprodutor será ponderada na proporção dos desvios padrão dos rebanhos nos quais essas filhas foram criadas. O resultado é que as produções das filhas, oriundas de rebanhos com maiores variâncias, influenciarão mais a avaliação de reprodutores do que a de filhas oriundas de rebanhos com menores variâncias. No caso de as filhas dos diversos reprodutores serem distribuídas, uniformemente, entre os rebanhos com variação diferente, nenhum vício seria observado. Entretanto, se as estimativas de herdabilidades também diferissem entre os rebanhos, a acurácia das avaliações genéticas dos animais seria reduzida, em razão de não serem consideradas essas diferenças (TORRES, 1998).

Veerkamp e Goddard (1998) estimaram a matriz de correlação genética e ambiental entre produções de leite, de gordura e de proteína durante os dez primeiros controles leiteiros de animais da raça Holandesa como características distintas, classificados em quatro classes de níveis de produção, totalizando 120 características diferentes. Posteriormente,

investigaram modelos alternativos para descrever a estrutura de (co)variâncias entre as características com um reduzido número de parâmetros usando a função de covariância em duas dimensões, estágio da lactação e nível de produção. Para a mesma característica no mesmo nível de produção as correlações genéticas entre controles variaram de 0,59 a 1,00, enquanto que as correlações ambientais oscilaram de 0,17 a 0,48. As correlações genéticas entre produções de leite e gordura, produção de leite e proteína e produção de gordura e proteína foram iguais a 0,38; 0,83 e 0,59, respectivamente. As correlações genéticas entre níveis de produção variaram de 0,79 a 0,97.

Utilizando modelos de regressão aleatória para descrever as covariâncias genéticas entre consumo de matéria seca, peso vivo e produção de leite, Veerkamp e Thompson (1999) estudaram as primeiras quinze semanas de lactação de 628 novilhas, com o objetivo de investigar estratégias para seleção em relação ao balanço energético dos animais. Estimativas de correlações genéticas entre a primeira e a décima quinta semana de lactação foram iguais a 0,62; 0,24 e 0,79 para as produções de leite, consumo de matéria seca e peso vivo, respectivamente. O consumo de matéria seca foi negativamente correlacionado com a produção de leite na primeira semana, porém foram positivamente correlacionados na décima quinta semana de lactação. A implicação deste resultado é que quando a seleção é baseada em uma combinação linear da produção de leite, consumo de

matéria seca e peso vivo, é importante considerar quando cada característica é mensurada durante a lactação.

Com o intuito de estudar a viabilidade da utilização de modelos de regressão aleatória para a descrição da variabilidade genética presente na produção de leite no dia do controle, o objetivo deste estudo consistiu em verificar a importância da inclusão do nível de produção dos rebanhos na formação de grupos contemporâneos, em modelos de avaliação genética dos reprodutores por meio de regressão aleatória.

2 MATERIAL E MÉTODOS

Os dados utilizados são provenientes de controles individuais de produção de leite no dia do controle, coletados mensalmente, pelo Serviço de Controle Leiteiro da Associação dos Criadores de Gado Holandês de Minas Gerais.

Os dados foram editados de forma que cada animal possuísse no mínimo três controles e que classe de rebanho e mês de controle apresentasse no mínimo quatro observações. Em cada controle, foi exigido, também, que os reprodutores possuíssem no mínimo duas progênies e, ainda, que cada classe rebanho-ano apresentasse no mínimo seis observações. O arquivo analisado continha 68 523 registros de controles individuais de produção de leite no dia do controle, provenientes de primeiras lactações pertencentes a 8 536 vacas, filhas de 537 reprodutores, distribuídas em 266 rebanhos, com parições nos anos de 1996 a

2001. As composições genéticas utilizadas neste estudo foram: animais puros de origem (PO) e animais de composição genética igual ou superior a 31/32 (PC = puros por cruza) Holandês. Os registros de controles mensais de produção de leite consideraram somente vacas ordenhadas duas vezes e três vezes ao dia. Os meses de parto das vacas foram agrupados em quatro estações, definidas como: estação 1, que corresponde aos meses de janeiro a março; estação 2, que corresponde aos meses de abril a junho; estação 3, correspondendo aos meses de julho a setembro e estação 4, que corresponde aos meses de outubro a dezembro. As estações 1 e 4 correspondem aos meses de maior incidência de chuvas e maiores temperaturas e umidade relativa do ar, enquanto que as estações 2 e 3 correspondem aos meses de menor incidência de chuvas e menores temperaturas e umidade relativa do ar.

As idades das vacas ao parto foram agrupadas em quatro classes definidas como: de 18 a 24 meses; maiores que 24 e menores que 26 meses; maiores que 26 e menores que 31 meses e maiores que 31 meses até 48 meses. As classes de idade da vaca ao parto foram distribuídas dentro das estações de parto.

Foi gerado, também, um arquivo de “pedigree”, utilizado em todas as análises, contendo a identificação de animal, pai e mãe, resultando em 17 664 indivíduos diferentes identificados na matriz de parentesco.

Com o objetivo de verificar a presença de heterogeneidade de variâncias entre rebanhos com produtividades diferentes e avaliar o seu efeito sobre a avaliação genética dos reprodutores por meio de modelos de regressão aleatória, classificaram-se os rebanhos em dois níveis de produção. Para classificar os rebanhos em dois níveis de produção, obteve-se, primeiramente, por meio de análise de componentes principais, combinações lineares da produção de leite para primeiro, segundo e terceiro controles em conjunto; para o quarto, quinto e sexto controles e, também, para o sétimo, oitavo e nono controles. O décimo controle não foi considerado na obtenção dos componentes principais devido ao menor número de registros. Em cada análise de componente principal foi considerada somente a primeira combinação linear. Posteriormente, de posse das três novas variáveis (componentes principais), obtiveram-se desvios padrão para cada rebanho, os quais foram classificados em dois níveis de produção, dispostos da seguinte forma: a) alto nível de produção, formado por rebanhos que apresentavam desvios padrão fenotípicos maiores que 5,85 kg; b) baixo nível de produção: referente aos demais rebanhos. Posteriormente, se formaram os grupos de contemporâneos de idade-estação de parto e níveis de produção. Assim, os valores genéticos dos reprodutores seriam preditos como desvios de uma curva de lactação média para cada classe de idade-estação-nível de produção. Estas mesmas predições foram comparadas com valores genéticos

preditos por um modelo que considerou o grupo de contemporâneo como idade-estação de parto. Do total de 537 reprodutores avaliados, 279 deles possuíam progênes em ambos níveis de produção. As distribuições dos controles leiteiros nos níveis de alta e baixa produção são exibidas nas Tabelas 2 e 3, respectivamente.

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

$$a_i^* = u + \frac{v - u}{a_{\max.} - a_{\min.}} (a_i - a_{\min.})$$

onde a_i^* é o número de dias em lactação padronizado; e u e v são definidos como $u = -1$ e $v = 1$; $a_{\max.}$ e $a_{\min.}$ são o maior e menor DIM mensurado, respectivamente.

O j -ésimo polinômio ortogonalizado de Legendre, P_j , é dado pela expressão:

$$\phi_j(a_i^*) = \frac{1}{2^j} \sqrt{\frac{2j+1}{2}} \sum_{m=0}^{[j/2]} (-1)^m \binom{j}{m} \binom{2j-2m}{j} a_i^{*j-2m}$$

em que $[j/2]$ indica que os valores da fração são arredondados para baixo, assumindo o valor inteiro mais próximo e DIMP os dias em lactação (DIM) padronizados.

O modelo de regressão aleatória utilizado para obtenção das soluções fixas e aleatórias por meio da função polinomial de Legendre, cujo grau do polinômio foi igual a segunda ordem ($k=3$) é descrito como:

$$y_{ijklmno} = RDC_i + GG_j + O_k + \sum_{m=0}^2 \beta_{(1)m} \phi_{nom} + \sum_{m=0}^2 \alpha_{nm} \phi_{nom} + \sum_{m=0}^2 \tau_{nm} \phi_{nom} + e_{ijklmno}$$

em que $y_{ijklmno}$ = é o registro o , da vaca n , realizado no dia em produção m , na subclasse idade-estação de parto-nível de produção l ; RDC_i é o efeito do rebanho-data i em que o controle foi realizado; GG_j = efeito da composição genética j ; O é o efeito do número de ordenhas diárias k ; $\beta_{(l)m}$ são os coeficientes de regressão obtidos em cada classe de idade-estação de parto-nível de produção, α_{nm} e τ_{nm} coeficientes de regressão aleatória, relacionados com os efeitos genético aditivo e de ambiente permanente, respectivamente, f_{no} refere-se ao número de dias em lactação (DIM) padronizados descritos pela função polinomial de Legendre de segunda ordem e $e_{ijklmno}$ é o erro aleatório associado a cada observação, $e_{ijklmno} \sim NID(0, \sigma_e^2)$.

Os valores genéticos preditos para a produção de leite acumulada aos 305 dias dos reprodutores foram obtidos como:

$$VG_{305} = \sum_{DIM=5}^{305} D_k = \left(\sum_{DIM=5}^{305} \phi_{no1} \right) \hat{a}_0 + \left(\sum_{DIM=5}^{305} \phi_{no2} \right) \hat{a}_1 + \left(\sum_{DIM=5}^{305} \phi_{no3} \right) \hat{a}_2$$

onde ϕ_{no1} ; ϕ_{no2} e ϕ_{no3} referem-se aos número de dias em produção padronizados e descritos pelos primeiro, segundo e terceiro polinômios de Legendre, respectivamente; $\hat{a}_0$, $\hat{a}_1$ e $\hat{a}_2$ são as soluções associadas com os três primeiros polinômios de Legendre, obtidas para o efeito genético aditivo por meio do modelo de regressão aleatória.

Os componentes de variância e covariância, necessários à estimação dos parâmetros genéticos e dos valores genéticos

dos animais, para a característica em estudo e em todas as análises, foram obtidos pelo método da Máxima Verossimilhança Restrita (REML), utilizando-se o sistema REMLF90 (MISZTAL, 1999) e as soluções dos coeficientes de regressão aleatória obtidas para cada animal, pelo sistema BLUPF90, ambos sistemas compilados para o sistema operacional LINUX.

3 RESULTADOS E DISCUSSÃO

O número de observações, médias com respectivos desvios padrão, valores mínimos e máximos para a produção de leite, podem ser observados na Tabela 1, respectivamente.

Os coeficientes de regressão aleatória estimados pelo modelo de regressão aleatória que utilizaram os polinômios ortogonais de Legendre de segundo grau, podem ser interpretados da seguinte forma: o primeiro parâmetro da função é uma constante, assumindo que os desvios de lactações individuais de cada animal em relação a uma curva média de lactação são constantes ao longo de toda a trajetória da lactação. O segundo parâmetro representa a fase ascendente da lactação, enquanto que o terceiro parâmetro determina a fase descendente da lactação. A seleção com base em um desses parâmetros com o objetivo de modificar a curva de lactação dos animais promovem respostas que não podem ser negligenciadas pelos outros parâmetros. A seleção para o aumento da taxa de produção de leite até o pico provoca indiretamente a seleção para animais com quedas mais bruscas da produção após o pico de produção, ou seja, seleciona animais de menor persistência.

Tabela 1 – Distribuição dos controles leiteiros com relação à produção de leite (cl1 a cl10), quantidade de informação (N), médias para a produção de leite ($\bar{X}$) e respectivos desvios padrão (DP), valores mínimos e máximos para a produção de leite, em kg, em cada controle leiteiro.

Controle	N	Média	DP	Mínimo	Máximo
cl1	8219	22,70968	5,8993	2,8	57,20
cl2	8354	25,09816	6,1686	4,4	53,00
cl3	8451	25,18124	6,3397	5,4	48,00
cl4	8045	24,86034	6,5071	3,0	52,00
cl5	7731	24,33006	6,7674	3,3	57,00
cl6	7466	23,49007	6,8279	2,4	62,80
cl7	7112	22,49215	6,8169	4,0	71,00
cl8	6426	21,57029	6,8529	2,6	78,60
cl9	4796	20,29138	6,7449	4,8	79,80
cl10	1923	19,43036	6,6021	4,4	69,40

Valores genéticos para a produção de leite acumulada aos 305 dias de produção para os reprodutores, preditos por meio de regressão aleatória utilizando a função polinomial de Legendre de segundo grau ($k=3$), foram obtidos quando suas progênes foram comparadas em grupos contemporâneos formados por classes de idade-estação de parto. Estas predições foram comparadas com os valores genéticos preditos quando o grupo contemporâneo além da idade-estação, considerou também o nível de produção dos rebanhos.

Os coeficientes de regressão aleatória estimados quando não foi considerado no modelo o grupo contemporâneo idade-

estação-nível de produção (Tabela 2) foram bem similares aos valores estimados quando se considerou o nível de produção (Tabela 3). Embora a função de verossimilhança tenha apresentado menor valor (240;430;8154 versus 240 475, 6325), quando se considerou o nível de produção, a estimativa da variância residual apresentou o mesmo valor quando o modelo incluiu o nível de produção (7,056 kg²).

Quando os animais são comparados na mesma classe de idade-estação-nível de produção de parto, os valores genéticos preditos para todos os animais foram menores quando comparados aos valores preditos quando os animais são comparados na mesma idade-estação de parto.

Tabela 2 – Coeficientes de regressão aleatória (diagonal e acima da diagonal) e correlações entre os coeficientes (abaixo da diagonal), estimados para o efeito genético aditivo (G) e para o efeito de ambiente permanente (P) por meio do modelo de regressão aleatória que não considerou o nível de produção na formação do grupo contemporâneo

Polinômio de segundo grau		a0	a1	a2
G	a0	11,3400	2,0240	-0,6339
	a1	0,5542	1,1760	-0,2839
	a2	-0,4888	-0,6798	0,1483
P	p0	19,8500	0,5656	-1,8090
	p1	0,0563	5,0740	0,2304
	p2	-0,2428	0,0611	2,7970

Tabela 3 – Coeficientes de regressão aleatória (diagonal e acima da diagonal) e correlações entre os coeficientes (abaixo da diagonal), estimados para o efeito genético aditivo (G) e para o efeito de ambiente permanente (P) por meio do modelo de regressão aleatória que considerou o nível de produção na formação do grupo contemporâneo

Polinômio de segundo grau		a0	a1	a2	
(k=3)	G	a0	11,1800	1,8440	-0,6671
		a1	0,5356	1,0600	-0,2797
		a2	-0,5032	-0,68519	0,1572
P		p0		p1	P2
		p0	19,9800	0,7237	-1,7570
		p1	0,0713	5,1430	0,2467
		p2	-0,2370	0,0656	2,7490

Quando se considera somente os 279 reprodutores que possuem filhas em comum em ambos níveis de produção, ao selecionar 10; 20; 30; 40; 50 e 100% dos melhores reprodutores (Tabela 4) com relação aos valores genéticos preditos para a produção de leite acumulada aos 305 dias de lactação,

quando o nível de produção não foi considerado na formação do grupo de contemporâneo e que apresentaram filhas em ambos níveis de produção, verifica-se o mesmo comportamento, ou seja, os valores genéticos preditos pelo modelo que considerou o nível de produção dos rebanhos

apresentaram menores valores médios, com maior amplitude entre o valor mínimo e máximo. A amplitude dos menores valores genéticos preditos foi maior quando se considerou o nível de produção dos rebanhos na formação do grupo de contemporâneo no modelo de avaliação genética, em comparação com a situação onde o nível de produção foi desconsiderado. Com relação à amplitude dos maiores valores genéticos preditos, observa-se pouca alteração. Este resultado sugere que ao considerar o nível de produção dos rebanhos, houve maior amplitude entre os valores genéticos preditos dos reprodutores. Assim, não considerar diferenças provenientes dos níveis de produção dos rebanhos no modelo de avaliação genética poderia resultar em superponderação da informação provenientes de reprodutores cujas filhas se concentrem mais no nível de alta produção, comparada com aqueles reprodutores dos quais as filhas se concentram em rebanhos de menor nível de produção. Resultados similares foram observados por Veerkamp e Goddard (1998) ao encontrarem correlações genéticas entre as produções de leite, gordura e proteína, distribuídas em quatro níveis de produção, variando de 0,79 a 0,97. Por outro lado, a correlação de produto-momento (correlação de Pearson) e de ordem (correlação de Spearman) entre os valores genéticos preditos para os reprodutores em ambos os níveis de produção e em cada percentagem de seleção dos melhores animais, verifica-se que todos os valores foram próximos da unidade,

indicando que na prática a inclusão do efeito do nível de produção dos rebanhos na formação dos grupos contemporâneos é irrelevante, pois a classificação dos animais quanto ao seu mérito genético para a produção de leite seria praticamente a mesma em ambas as situações.

As soluções obtidas para cada animal por meio do modelo de regressão aleatória, por si próprias não são suficientes para ordenar ou selecionar os animais, porém funções destas soluções provêm uma informação eficiente para a decisão no ordenamento e seleção dos animais avaliados. Estas soluções são obtidas como desvios de uma curva média obtida para cada grupo de contemporâneo. Desta forma, ao prever os valores genéticos dos animais (por meio destas soluções) quando o grupo contemporâneo foi formado por classes de idade-estação de parto e o nível de produção dos rebanhos, esperou-se que as soluções obtidas para cada animal refletisse também diferenças entre níveis de produção dos rebanhos, uma vez que reprodutores cujas filhas se concentrasse em rebanhos de maior produção teriam seus valores genéticos preditos superestimados, em relação ao reprodutores cujas a maior parte das filhas se concentrasse em rebanhos de menor produção. Ao comparar estas predições com os valores genéticos preditos quando o grupo de contemporâneo foi formado somente por idade-estação de parto, a correlação de Pearson (ordem) foi próximo de um, indicando que, independentemente

da situação, os animais seriam classificados da mesma forma. Embora este resultado indique que nenhuma diferença é esperada no ordenamento dos melhores indivíduos por meio das predições obtidas em ambas situações, uma implicação prática é que os animais ordenados como primeiros em uma

situação poderiam, não necessariamente, ser os primeiros na outra situação, tendo reflexo no preço de sêmem de um determinado reprodutor. Foi verificado neste estudo, que até o décimo terceiro colocado não houve diferenças nas classificações em ambas situações.

Tabela 4 – Estimativas de médias dos valores genéticos preditos para a produção dos reprodutores por meio de modelo de regressão aleatória quando o grupo contemporâneo incluiu os efeito de idade- estação (VG1) e idade-estação-nível de produção (VG2), valores máximos e mínimos, coeficientes de correlação de Pearson e de Spearman, para diferentes porcentagens de seleção dos reprodutores

Porcentagem Seleccionada				Valore genéticos preditos		Correlação de	
		N	$\bar{X}$	Mínimo	Máximo	Pearson	Spearman
10	VG1	27	923,217	675,901	1447	0,995	0,983
	VG2	27	914,423	656,365	1449		
20	VG1	55	747,322	478,496	1447	0,983	0,980
	VG2	55	734,858	456,804	1449		
30	VG1	83	627,378	312,285	1447	0,995	0,988
	VG2	83	616,371	292,54	1449		
40	VG1	111	534,780	193,059	1447	0,995	0,991
	VG2	111	524,537	149,648	1449		
50	VG1	139	454,080	70,611	1447	0,991	0,993
	VG2	139	446,003	46,995	1449		
100	VG1	279	75,596	-1567	1447	0,999	0,998
	VG2	279	70,635	-1578	1449		

4 CONCLUSÃO

Com base no objetivo deste estudo e nos resultados obtidos, pode-se afirmar que a inclusão do efeito de nível de produção entre rebanhos, na formação de grupos de contemporâneos para a predição dos valores genéticos dos reprodutores, por meio de modelos de regressão aleatória, não é relevante e conduzem aos mesmos resultados quando desprezado.

REFERÊNCIAS

- BOLDMAN, K. G.; FREEMAN, A. E. Adjustment for heterogeneity of variance by herd production level in dairy cow and sire evaluation. *Journal of Dairy Science*, v.73, n.2, p.503-512, 1990.
- COSTA, C. N. *Genetic relationships for milk and fat yields between Brazilian and United States Holstein cattle populations*. 1998. 175 p. Thesis (Doctor of Philosophy) - Cornell University, Ithaca, 1998.
- DODENHOFF, J.; SWALVE, H. H. Heterogeneity of variances across regions of northern Germany and adjustment in genetic evaluation. *Livestock Production Science*, v. 53, n.1, p.225-236, 1998.
- DONG, M. C.; MAO, I. L. Heterogeneity of (co)variance and heritability in different levels of intra-herd milk production variance and of herd average. *Journal of Animal Science*, v.73, n.3, p.843-851, 1990.
- ENGEL, B.; MEUWISSEN, T.; DE JONG, G., BUIST, W. Prediction of breeding values with a mixed model with heterogeneous variances for large-scale dairy data. *Journal of Agricultural, Biological and Environmental Statistics*, v. 4, n.2, p.185-201, 1999.
- HENDERSON JUNIOR, C. R. Analysis of covariance in the mixed model: higher level, nonhomogeneous, and random regressions. *Biometrics*, v.38, p.623-640, 1982.
- HILL, W. G.; EDWARDS, M. R.; AHMED, M. K. A.; THOMPSON, R. Heritability of milk yield and composition at different levels and variability of production. *Animal Production*, v.36, n.1, p. 59-68, 1983.
- JAMROZIK, J.; SCHAEFFER, L. R.; DEKKERS, J. C. M. Genetic evaluation of dairy cattle using test day yields and random regression model. *Journal of Dairy Science*, v. 80, n. 6, p. 1217-1226, 1997.
- KIRKPATRICK, M.; LOFSVOLD, D.; BULMER, M. Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics*, v. 24, n. 3, p. 979-993, 1990.
- MEYER, K. Estimates of variance due to sire x herd interactions and environmental covariances between paternal half-sibs for first lactation dairy production. *Livestock Production Science*, v.17,n. 1, p.95-115, 1987.

- MEYER, K.; HILL, W. G. Estimation of genetic and phenotypic covariance functions for longitudinal or "repeated" records by restricted maximum likelihood. *Livestock Production Science*, v. 47, n. 3, p. 185-200, 1997.
- MISZTAL, I. *REMLF90 Manual*. 1999. Disponível em: <ftp://nce.ads.uga.edu/pub/ignacy/blupf90/>. Acesso em: jan. 2001.
- POOL, M.H.; MEUWISSEN, T.H.E. Prediction of daily milk yields from a limited number of test days using test day model *Journal of Dairy Science*, v. 82, n. 7, p. 1555-1564, 1999.
- SCHAEFFER, L. R.; DEKKERS, J. C. M. Random regression in animal models for test day production in dairy cattle. In: WORLD CONGRESS GENETIC APPLIED LIVESTOCK PRODUCTION, 5., 1994, Guelph. *Proceedings...* Guelph, 1994. p. 443-446.
- STANTON, T. L.; BLAKE, R. W.; QUAAS, R. L.; VAN VLECK, L. D.; CARABAÑO, M. J. Genotype by environment interaction for Holstein milk yield in Colombia, México and Porto Rico. *Journal of Dairy Science*, v.74, n.5, p.1700-1714, 1991.
- TORRES, R. A. *Efeito da heterogeneidade de variância na avaliação genética de bovinos da raça Holandesa no Brasil*. 1998. 124 p. Tese (Doutorado em Ciência Animal) - Universidade Federal de Minas Gerais, Belo Horizonte, 1998.
- VEERKAMP, R. F.; GODDARD, M. E. Covariance functions across herd production levels for test day records on milk, fat and protein yields *Journal of Dairy Science*, v. 81, n. 6, p. 1690-1701, 1998.
- ; THOMPSON, R. A covariance function for feed intake, live weight, and milk yield estimated using a random regression model. *Journal of Dairy Science*, v. 82, n.6, p.1565-1573, 1999.