



## Comparação de modelos de regressão aleatória para estimação de parâmetros genéticos em caprinos leiteiros

José Lindenberg Rocha Sarmiento<sup>1\*</sup>, Lucia Galvão de Albuquerque<sup>2,4</sup>, Robledo de Almeida Torres<sup>3,4</sup>, Marcelo Teixeira Rodrigues<sup>3</sup>, Paulo Sávio Lopes<sup>3,4</sup>, João Cruz Reis Filho<sup>1</sup>

<sup>1</sup> Pós-graduação em Genética e Melhoramento Animal/UFV. Bolsista do CNPq.

<sup>2</sup> FCAV/UNESP, CEP: 14884-900, Jaboticabal-SP.

<sup>3</sup> Departamento de Zootecnia/UFV, 36571-000, Viçosa-MG.

<sup>4</sup> Bolsista do CNPq.

**RESUMO** - Objetivou-se avaliar a melhor modelagem para as variâncias genética aditiva, de ambiente permanente e residual da produção de leite no dia do controle (PLDC) de caprinos. Utilizaram-se modelos de regressão aleatória sobre polinômios ortogonais de Legendre com diferentes ordens de ajuste e variância residual heterogênea. Consideraram-se como efeitos fixos os efeitos de grupo de contemporâneos, a idade da cabra ao parto (co-variável) e a regressão fixa da PLDC sobre polinômios de Legendre, para modelar a trajetória média da população; e, como efeitos aleatórios, os efeitos genético aditivo e de ambiente permanente. O modelo com quatro classes de variâncias residuais foi o que proporcionou melhor ajuste. Os valores do logaritmo da função de verossimilhança, de AIC e BIC apontaram para seleção de modelos com ordens mais altas (cinco para o efeito genético e sete para o efeito de ambiente permanente). Entretanto, os autovalores associados às matrizes de co-variâncias entre os coeficientes de regressão indicaram a possibilidade de redução da dimensionalidade. As altas ordens de ajuste proporcionaram estimativas de variâncias genéticas e correlações genéticas e de ambiente permanente que não condizem com o fenômeno biológico estudado. O modelo de quinta ordem para a variância genética aditiva e de sétima ordem para o ambiente permanente foi indicado. Entretanto, um modelo mais parcimonioso, de quarta ordem para o efeito genético aditivo e de sexta ordem para o efeito de ambiente permanente, foi suficiente para ajustar as variâncias nos dados.

Palavras-chave: componentes de variância, dia do controle, modelo animal, polinômios de Legendre, produção de leite

## Comparison of random regression models for the estimation of genetic parameters in dairy goats

**ABSTRACT** - Random regression and Legendre polynomial (LP) of different orders were used for modeling the genetic, permanent environmental and residual variances of test day milk yield in dairy goats. The models included the fixed effects of contemporary group, age of dam at kidding as a covariate and the fixed regression of LP for the average lactation curve and the additive genetic, permanent environmental and residual as random effects. According to the values of the logarithm of the likelihood function, AIC and BIC using higher orders of LP (fifth order for the genetic effect and seventh order for the permanent environmental effect) improved the fitting of the models. The model with four classes of residual variances provided the best fit. The eigenvalues of the (co)variances matrix among the regression coefficients suggested the possibility of reducing the dimension of the models. The estimates of genetic variances and genetic and permanent environmental correlations for test day milk yields obtained from polynomials of higher orders are not biologically expected. The LP of fifth order for the additive genetic and seventh order for the permanent environmental effects was the best fitted model. However, a LP of fourth order for the additive genetic and of sixth order for the permanent environmental effects may be considered as a more parsimonious model for the estimation of variances of test day milk yield in dairy goats, by random regression.

Key Words: animal model, Legendre polynomials, milk yield, test day, variance components

### Introdução

A produção de leite apresenta caráter contínuo que, biologicamente, pode ser explicado por diferentes genes que se expressam no decorrer da lactação, causando alte-

rações fisiológicas e de produção. Entretanto, assume-se a descontinuidade dessa característica, ou seja, os dados de produção de leite medidos a cada controle durante a lactação têm sido utilizados apenas para calcular a produção total na lactação ou em determinados dias em lactação. Atualmente,

Este artigo foi recebido em 15/12/2006 e aprovado em 12/5/2008.

Correspondências devem ser enviadas para [sarmiento@ufpi.br](mailto:sarmiento@ufpi.br).

\*Endereço atual: Campus de Bom Jesus/UFPI, Bom Jesus-PI, CEP: 64.900-000.

a produção de leite no dia de controle (PLDC) tem sido empregada em substituição à produção total nas avaliações genéticas de caprinos leiteiros no Canadá (Sullivan & Wiggans, 2000), por meio de modelos de produção no dia de controle, os chamados “test day models”.

O surgimento dos modelos de regressão aleatória (Henderson Jr., 1982) possibilitou assumir a continuidade de características que se repetem ao longo do tempo, como a PLDC, o que pode aumentar a acurácia e a possibilidade de seleção e melhoramento genético. Em comparação aos tradicionais modelos animais (repetibilidade e multicausalidade), que fornecem previsões para determinados pontos da curva ou dias em lactação, os modelos de regressão aleatória permitem a estimação de parâmetros genéticos e a previsão de valores genéticos para quaisquer dias em lactação. Ao ajustar um modelo de regressão aleatória, assume-se implicitamente determinada estrutura de covariâncias entre as produções nos controles, imposta pelo modelo de regressão aleatória escolhido. Assim, consideram-se as correlações na análise e assume-se determinada estrutura para as correlações. Além disso, esses modelos possuem todas as vantagens atribuídas ao uso da PLDC, discutidas por Sarmiento et al. (2006).

Schaeffer (2004) sugere que os modelos de regressão aleatória, assim como em vacas leiteiras, devem ser aplicados também em cabras leiteiras com os devidos ajustes de acordo com as particularidades intrínsecas da espécie, como intervalos de tempo entre as medidas, definição de subclasses para a curva fixa e da ordem de ajuste da regressão aleatória.

Neste sentido, objetivou-se com este trabalho comparar modelos com diferentes ordens de ajuste para a regressão aleatória genética aditiva e de ambiente permanente, por meio de funções polinomiais de Legendre com variância residual homogênea ou heterogênea para buscar o modelo mais adequado para descrever as mudanças nas variâncias associadas à produção de leite de cabras.

## Material e Métodos

Analisaram-se 9.374 registros semanais de PLDC provenientes de 302 primeiras lactações de cabras da raça Alpina, pertencentes ao rebanho do Departamento de Zootecnia da Universidade Federal de Viçosa (Viçosa, Minas Gerais). Os controles foram registrados semanalmente e obtidos de ordenha mecânica, realizada duas vezes ao dia. Os animais foram mantidos em baias coletivas sob o sistema de estabulação livre. Os dados foram coletados entre os anos de 1997 e 2004. O banco de dados inicial era

composto de 17.482 registros de PLDC, no entanto, foram considerados os registros de controles da 2ª até a 39ª semana de lactação (correspondendo a até 270 dias em lactação) e excluídos da análise animais com produções superiores ou inferiores a três desvios-padrão da média da semana da lactação – produções que ocorreram em grupos contemporâneos (ano-quizena do controle) com menos de cinco cabras e lactações com menos de seis controles (Tabela 1).

A produção de leite no dia do controle (PLDC) foi analisada pelo modelo animal, unicaracterístico, de regressão aleatória descrito a seguir:

$$y_{ij} = F + \sum_{m=0}^{k_b-1} b_m f_m(t_{ij}) + \sum_{m=0}^{k_a-1} a_{im} f_m(t_{ij}) + \sum_{m=0}^{k_g-1} g_{im} f_m(t_{ij}) + e_{ij}$$

em que:  $y_{ij}$  = produção de leite no controle  $j$  da cabra  $i$ ;  $F$  = conjunto de efeitos fixos, constituído do grupo contemporâneo (138 subclasses) e da co-variável idade da cabra ao parto, efeitos linear e quadrático;  $b_m$  = coeficiente de regressão  $m$  da PLDC sobre o polinômio de Legendre para modelar a curva média da população;  $a_{im}$  e  $g_{im}$  =  $m$  coeficientes de regressão genético aditivo e de ambiente permanente, respectivamente, para a cabra  $i$ ;  $k_b$ ,  $k_a$  e  $k_g$  = ordens dos polinômios de Legendre;  $(t_{ij})$  = semana de lactação  $j$  da cabra  $i$  padronizada para o intervalo de -1 a 1;  $f_m(t_{ij})$  = função polinomial de Legendre  $m$  a ser avaliada para  $(t_{ij})$ ; e  $e_{ij}$  = efeito de ambiente temporário afetando  $y_{ij}$ .

A regressão que gerou a trajetória média da população (regressão fixa) foi modelada por meio de polinômios ortogonais de Legendre da semana de lactação padronizada

Tabela 1 - Número de registros (NR), média e desvio-padrão (DP) para produção de leite na semana do controle da primeira lactação de cabras da raça Alpina

Semana	NR	Média (kg)	DP	Semana	NR	Média (kg)	DP
2	160	2,41	0,88	21	270	2,24	0,78
3	183	2,57	0,97	22	241	2,21	0,80
4	219	2,54	0,92	23	198	2,23	0,76
5	207	2,61	0,87	24	195	2,15	0,82
6	225	2,70	0,89	25	210	2,12	0,77
7	260	3,00	1,15	26	194	2,06	0,74
8	318	3,12	1,15	27	189	2,11	0,76
9	334	3,10	1,13	28	181	1,96	0,71
10	380	3,04	1,13	29	156	1,90	0,75
11	386	2,98	1,06	30	169	1,87	0,70
12	413	2,88	1,07	31	180	1,92	0,79
13	407	2,85	1,07	32	162	1,93	0,81
14	464	2,80	1,04	33	153	1,98	0,83
15	458	2,76	1,00	34	147	1,92	0,77
16	375	2,51	0,90	35	155	1,89	0,76
17	371	2,42	0,88	36	129	1,77	0,80
18	345	2,44	0,88	37	136	1,79	0,80
19	315	2,34	0,83	38	129	1,89	0,85
20	268	2,27	0,79	39	92	1,88	0,83

(-1 a 1) de ordem 3 ( $k_b = 3$ ), para todos os modelos analisados. As regressões aleatórias também foram modeladas por polinômios ortogonais de Legendre, variando à ordem de ajuste de terceira à quinta ( $k_a = 3, 4$  e 5) para o componente genético aditivo e de terceira à sétima ( $k_g = 3, \dots, 7$ ) para o componente de ambiente permanente.

A variância residual foi considerada, inicialmente, homogênea durante toda a lactação. Posteriormente, para considerar heterogeneidade de variância em classes, assumiram-se duas (da 2<sup>a</sup> à 15<sup>a</sup> e da 16<sup>a</sup> à 39<sup>a</sup> semana de lactação), três (da 2<sup>a</sup> à 6<sup>a</sup>, da 7<sup>a</sup> à 15<sup>a</sup> e da 16<sup>a</sup> à 39<sup>a</sup> semana de lactação), quatro (da 2<sup>a</sup> à 6<sup>a</sup>, da 7<sup>a</sup> à 10<sup>a</sup>, da 11<sup>a</sup> à 15<sup>a</sup> e da 16<sup>a</sup> à 39<sup>a</sup> semana de lactação), cinco (da 2<sup>a</sup> à 6<sup>a</sup>, da 7<sup>a</sup> à 10<sup>a</sup>, da 11<sup>a</sup> à 15<sup>a</sup>, da 16<sup>a</sup> à 26<sup>a</sup> e da 27<sup>a</sup> à 39<sup>a</sup> semana de lactação) e seis (da 2<sup>a</sup> à 6<sup>a</sup>, da 7<sup>a</sup> à 10<sup>a</sup>, da 11<sup>a</sup> à 15<sup>a</sup>, da 16<sup>a</sup> à 26<sup>a</sup>, da 27<sup>a</sup> à 32<sup>a</sup> e da 33<sup>a</sup> à 39<sup>a</sup> semana de lactação) classes de variâncias residuais. Nesses casos, as variâncias foram assumidas como heterogêneas entre os grupos semanais e homogêneas dentro de cada grupo.

Os modelos estudados serão referenciados como  $Ak_a Ck_g Re$ , de modo que A, C e R representam os efeitos genéticos aditivos, de ambiente permanente e residual;  $k_a$  e  $k_g$  como definidos anteriormente; e  $e$  representa o número de classes ajustadas para variância residual.

As co-variâncias entre os coeficientes de regressão aleatória genéticos e de ambiente permanente foram estimadas pelo método da máxima verossimilhança restrita (REML) utilizando-se o programa DXMRR do DFREML (Meyer, 1998).

As comparações entre os modelos foram feitas pelo Critério de Informação de Akaike (AIC), pelo Critério de Informação Bayesiano de Schwarz (BIC), pelo logaritmo da função de verossimilhança (Log L) e pelo teste da razão de verossimilhança (TRV) a 1% de probabilidade. As estimativas de variância genética, de ambiente permanente e fenotípica e os parâmetros genéticos resultantes, obtidos pelos modelos estudados, foram analisados e comparados entre si e com as respectivas estimativas obtidas por meio de um modelo para PLDC ordinário, realizadas por Sarmiento et al. (2006) no mesmo conjunto de dados.

Para determinar a importância relativa dos autovalores da matriz de coeficientes de regressão aleatória referentes ao efeito genético aditivo na explicação da variação genética total, cada autovalor foi expresso em porcentagem da soma de todos os autovalores.

## Resultados e Discussão

Inicialmente, com o objetivo de avaliar o melhor ajuste para a variância residual, considerou-se a variância como homogênea e em classes heterogêneas (Tabela 2). Para

Tabela 2 - Número de parâmetros estimados (NP), logaritmo da função de máxima verossimilhança (Log L), critério de informação de Akaike (AIC), critério de informação Bayesiano (BIC) e teste da razão de verossimilhança (TRV) para os modelos utilizados na comparação do número de classes para a variância residual

Modelo	NP	Log L <sup>1</sup>	AIC <sup>1</sup>	BIC <sup>1</sup>	TRV
1 - R1	13	-122	234	209	(2-1) 46,8*
2 - R2	14	-98	189	171	(3-2) 122,0*
3 - R3	15	-37	69	59	(4-3) 67,6*
4 - R4	16	-3	4	0	(5-4) 5,6 <sup>ns</sup>
5 - R5	17	-1	0	4	(6-5) 1,2 <sup>ns</sup>
6 - R6	18	0	1	116	-

\* Significativo a 1% de probabilidade; ns: não-significativo.

<sup>1</sup> Valores expressos como desvio do melhor valor.

isso, manteve-se constante a ordem de ajuste para os efeitos genéticos aditivos e de ambiente permanente ( $k_a = k_g = 3$ ). Comparando os modelos homogêneo e heterogêneo com duas classes, constataram-se aumento no Log L e diminuição nos critérios AIC e BIC. Observou-se também que a diferença no Log L entre os dois modelos foi estatisticamente significativa ( $P < 0,01$ ) pelo TRV (Tabela 2). De acordo com os critérios utilizados para avaliar a qualidade de ajuste, o modelo considerando homogeneidade de variâncias residuais mostrou-se inadequado em comparação àquele com duas classes.

As variâncias residuais heterogêneas durante a lactação observadas neste trabalho têm sido relatadas por vários autores (Rekaya et al., 2000; El Faro & Albuquerque, 2003) e podem ser atribuídas a fatores como estágio de prenhez, condições do parto, duração do intervalo de lactações, entre outros, uma vez que estes fatores não são facilmente incorporados nos modelos de análise, em virtude da falta de informação sobre eles (Assis et al., 2006). Com isso, essa variação é direcionada para o resíduo, tornando necessário modelá-la.

O modelo com três classes (R3), quando ajustado, apresentou mudança de maior magnitude nos critérios utilizados, em comparação ao modelo R2, ou seja, diminuição no AIC e BIC e aumento no Log L altamente significativo ( $P < 0,01$ ) pelo TRV. Aumento no Log L foi verificado nos demais modelos ajustados, isto é, o aumento no número de classes (de R1 a R6) promoveu aumento no Log L, como esperado, uma vez que este critério normalmente indica o modelo com maior número de parâmetros a ser ajustado. Segundo esse critério, o modelo com seis classes deveria ser aplicado, entretanto, a partir do modelo R4, as mudanças no Log L foram de pequena magnitude (Tabela 2) e não-significativas pelo TRV ( $P > 0,01$ ), o que pode indicar que

modelos com quatro classes poderiam ser suficientes para se obter bom ajuste da variância residual.

Os valores de AIC diminuíram até o modelo R5 (Tabela 2) e tenderam a apresentar ligeiro aumento com o ajuste do modelo R6, sugerindo que a criação da sexta classe não melhorou o ajuste da variância residual. Esse resultado é consequência da penalidade que este critério impõe a modelos mais parametrizados. O comportamento observado com o AIC foi semelhante àquele obtido com o Log L, exceto pelo fato de este critério ter indicado o modelo R5 como o que melhor se ajustou à variância residual, embora a diferença em relação ao modelo R4 tenha sido desprezível.

As maiores mudanças em melhoria de ajuste nos valores do BIC (Tabela 2), que penaliza com maior rigor que o AIC, em razão do número de parâmetros, ocorreram até o modelo R4, enquanto o R5 e o R6 foram inferiores.

Os resultados dos critérios analisados conjuntamente sugerem que um modelo com quatro ou cinco classes é suficiente para modelar a variância residual. Na comparação da variância residual entre os modelos R1, R4 e R5 (Figura 1), verificou-se diferença em considerar um modelo homogêneo ou heterogêneo, principalmente na fase inicial da lactação, quando houve maior discrepância entre as estimativas obtidas para variância residual, o que pode influenciar na partição da variância total entre os demais componentes de variância a ser estimados. As estimativas obtidas pelos modelos R4 e R5 foram muito semelhantes (Figura 1) e sugerem que quatro classes (R4) foram suficientes para modelar a variância residual ao longo da curva de lactação para o rebanho em estudo.

A modelagem da regressão aleatória por meio de funções polinomiais de Legendre requer, inicialmente, a definição da ordem mais apropriada para cada efeito aleatório considerado no modelo de análise. Em síntese, na modelagem da variância genética aditiva (Tabela 3), observou-se melhoria nos critérios Log L ( $P < 0,01$ , pelo TRV) e AIC com a elevação da ordem de ajuste de três para cinco e essa melhoria foi sempre associada à maior ordem para o efeito de ambiente permanente. O critério BIC, mais rigoroso em virtude da parametrização, apontou para a exclusão do modelo de quinta ordem para este efeito. Na modelagem do componente de ambiente permanente, todos os critérios anteriores apontaram a sétima ordem como a que proporcionou melhor ajuste, o que reflete a necessidade de ordens mais elevadas para esse efeito.

A seleção de modelos de regressão aleatória utilizando critérios de comparação que os ordena quanto ao melhor ajuste de diferentes formas não é trivial e tem sido objetivo de pesquisa nos últimos anos (Druet et al., 2003; López-Romero & Carabaño, 2003; Araújo et al., 2006). Em estudo realizado por Ødegard et al. (2003), a escolha do modelo foi baseada na inspeção simultânea dos componentes de variância estimados, no teste da razão de verossimilhança, em critérios de informação baseados na função de verossimilhança, na variância residual e no quadrado médio do erro de predição. Neste estudo, não houve concordância entre os três critérios utilizados, isto é, o Log L e o AIC apontaram o modelo A5C7R4 como o melhor, entretanto, pelo BIC, o A4C7R4 deveria ser o escolhido.

Ressalta-se que, de acordo com Legarra et al. (2004), ajuste perfeito não é necessário, uma vez que esse ajuste

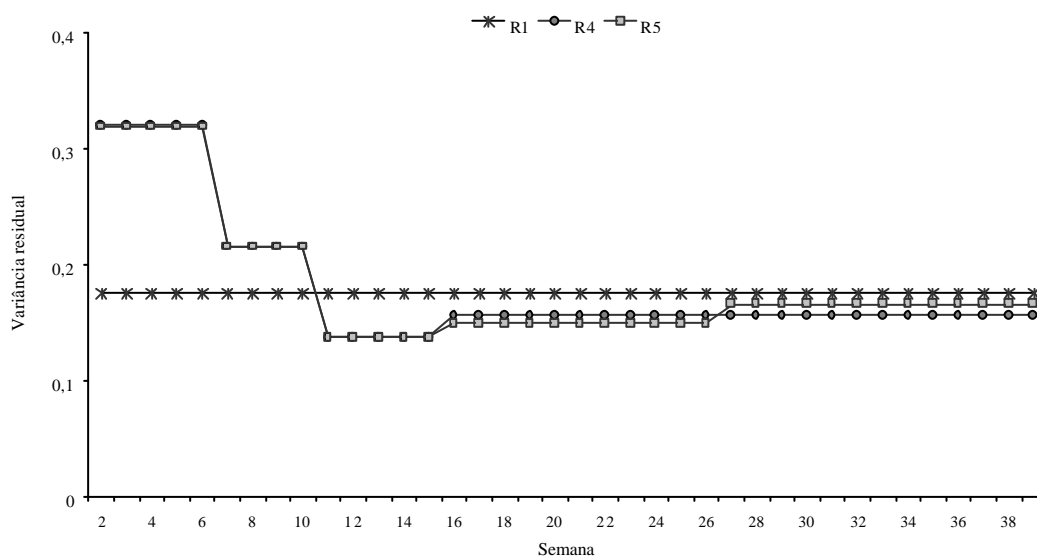


Figura 1 - Estimativas de variâncias residuais homogêneas e heterogêneas com uma, quatro e cinco classes, obtidas, respectivamente, pelos modelos R1, R4 e R5.

Tabela 3 - Número de parâmetros estimados (NP), logaritmo da função de máxima verossimilhança (Log L), critério de informação de Akaike (AIC), critério de informação Bayesiano (BIC) e teste da razão de verossimilhança (TRV) para os modelos com diversas ordens para os efeitos aleatórios

Modelo	NP	Log L <sup>1</sup>	AIC <sup>1</sup>	BIC <sup>1</sup>	TRV
1 - A3C3R4	16	-349	637	446	(2-1)318,6*
2 - A3C4R4	20	-190	326	164	(3-2)154,8*
3 - A3C5R4	25	-113	181	55	(4-3)36,9*
4 - A4C5R4	29	-94	153	54	(5-4)86,3*
5 - A4C6R4	35	-51	78	23	(6-5)86,8*
6 - A4C7R4	42	-8	5	0	(7-4)15,1*
7 - A5C5R4	34	-87	147	85	(8-7)83,1*
8 - A5C6R4	40	-45	76	57	(9-8)90,2*
9 - A5C7R4	47	0	0	30	-

\* Significativo a 1% de probabilidade.

<sup>1</sup> Valores expressos como desvio do melhor valor.

aumenta a demanda computacional (memória e capacidade de processamento) e a susceptibilidade a erros numéricos. Assim, modelos mais parcimoniosos devem ser preferidos. Entretanto, como mencionado, a seleção da ordem do polinômio a ser utilizada não é uma tarefa fácil e, para isso, os autovalores podem ser uma ferramenta útil (Foulley & Robert-Granié, 2002).

O primeiro autovalor associado à matriz dos coeficientes de regressão aleatória (Tabela 4), tanto dos efeitos genético aditivo e de ambiente permanente, foi responsável, em todos os modelos ajustados, por mais de 70% da variação nos dados. Houve mudança na proporção da variância explicada pelos autovalores com o aumento da ordem de ajuste das funções polinomiais.

Quando uma função de terceira ordem foi ajustada para o efeito genético aditivo, obteve-se um autovalor que não explicou nada da variação observada, ao passo que, quando uma função de ordem 4 foi ajustada, houve redistribuição da porção explicada entre os autovalores, ou seja, o terceiro passou a explicar 2,8% da variação total, embora o quarto não tenha explicado nada. Ao passar da quarta para a quinta ordem, o quarto e quinto autovalores explicaram 0% da variação, exceção feita para o quarto, que explicou apenas 0,2% quando o efeito de ambiente permanente foi modelado com polinômio de Legendre de ordem 7. Aumentando a ordem da função polinomial para o efeito genético aditivo de três para quatro, houve tendência de aumento na porção da variância explicada pelo segundo e terceiro autovalores. Esse aumento foi ligeiramente maior quando o efeito de ambiente permanente foi ajustado por uma função de sétima ordem. Entretanto, o aumento da ordem de 4 para 5 não proporcionou mudanças na proporção explicada pelos autovalores, como aconteceu com o aumento de 3 para 4,

Tabela 4 - Porcentagem da variância explicada pelos autovalores da matriz dos coeficientes de regressão aleatória genética aditiva e de ambiente permanente para modelos selecionados

Efeito	Coeficiente de regressão aleatória						
	1	2	3	4	5	6	7
A3C5R4							
A3	92,3	7,7	0	-	-	-	-
C5	71,5	15,5	7,6	3,7	1,7	-	-
A4C5R4							
A4	89,2	8,0	2,8	0	-	-	-
C5	72,4	16,3	7,0	2,6	1,7	-	-
A4C6R4							
A4	88,5	8,6	2,9	0	-	-	-
C6	71,4	15,8	7,1	3,4	2,0	0,4	-
A4C7R4							
A4	86,7	8,7	4,6	0	-	-	-
C7	70,3	15,9	7,3	3,5	2,2	0,8	0
A5C5R4							
A5	86,5	8,5	5,0	0	0	-	-
C5	73,6	16,3	6,7	2,7	0,7	-	-
A5C6R4							
A5	86,8	8,9	4,3	0	0	-	-
C6	72,4	16,0	6,8	3,3	1,2	0,3	-
A5C7R4							
A5	84,2	9,7	5,9	0,2	0	-	-
C7	71,1	15,4	7,3	3,1	2,6	0,5	0

portanto, esse aumento não contribuiu, segundo os autovalores, no ajuste da variância genética aditiva.

Para o efeito de ambiente permanente, quando modelado pelas funções de quinta e sexta ordem, todos os autovalores explicaram um percentual da variação, embora a porção da variância explicada tenha diminuído com o aumento da ordem. Para esse efeito, quando a ordem de ajuste da função foi 7, apenas o último autovalor, independentemente da ordem para o efeito genético aditivo, não foi importante, o que aponta para a exclusão desta ordem.

Pela análise dos autovalores, a dimensionalidade dos dois efeitos aleatórios poderia ser reduzida sem perda de informação, discordando do Log L e do critério AIC para os efeitos genético aditivo e ambiente permanente e concordando com o critério BIC para o efeito genético. Contudo, Legarra et al. (2004) alertaram para o fato de que a redução da dimensionalidade decorrente da eliminação de autovalores próximos de zero não é indicada em todos os casos, uma vez que, ao adotar este critério, pode resultar em modelagem simplista ou inadequada. Assim, optou-se por analisar os componentes de variância estimados pelos modelos e os parâmetros genéticos resultantes.

Como se esperava, os modelos ajustaram de forma semelhante à variância fenotípica e foram compatíveis com as obtidas em análises univariadas (Figura 2). Observaram-se, de modo geral, pequenas diferenças na variância

de ambiente permanente estimada pelos modelos estudados, sendo estas, no início, próximas ao pico e no final da curva de lactação. Na primeira semana, os modelos A4C5R4, A4C6R4 e A4C7R4 tenderam a estimar variâncias maiores; no pico, os modelos A4C7R4 e A5C7R4 estimaram variâncias superiores; e no final da curva de lactação, os modelos A4C5R4 e A5C5R4 tenderam a estimar valores ligeiramente superiores.

Para a variância genética aditiva, os modelos mais parametrizados (A5C5R4, A5C7R4, A5C6R4 e A4C7R4) estimaram variâncias genéticas mais elevadas na primeira semana de lactação, tendendo a decrescer até a sexta semana de lactação (Figura 2). Os modelos A4C6R4 e A4C5R4, no entanto, apresentaram comportamento contrário, com variâncias menores na primeira e aumentando até próximo à décima semana (em torno do pico de lactação), com tendência a decrescer a partir deste ponto.

Do pico à 22ª semana de lactação, as variâncias estimadas por todos os modelos decresceram nas mesmas proporções. A partir deste ponto, os modelos A4C7R4 e A5C7R4 tenderam a estimar valores mais altos para variância genética aditiva; as variâncias estimadas pelo modelo A5C7R4 foram de maior magnitude, exceto a partir da 36ª semana, quando os modelos A4C7R4, A4C6R4 e A4C5R4 estimaram variâncias repentinamente superiores (Figura 2), o que não parece ser biologicamente correto.

Estimativas de variâncias genéticas inconsistentes foram estimadas pelos seis modelos (Figura 2); os de ordem 4 apresentaram problemas no final e os de ordem 5, no início da curva de lactação. Esses problemas no início e no final das curvas podem estar relacionados à utilização de modelos

inadequados para os diferentes efeitos aleatórios (Van der Werf et al., 1998), que pode não ter ocorrido neste trabalho, em razão do menor número de observações no final da curva (Strabel & Misztal, 1999) ou às propriedades matemáticas da função escolhida (Kettunen et al., 1998).

As herdabilidades estimadas pelos modelos estudados apresentaram a mesma tendência descrita para as estimativas de variâncias genéticas (Figura 3), ou seja, os modelos A5C5R4, A5C7R4, A5C6R4 e A4C7R4 estimaram herdabilidades altas no início da lactação (0,50; 0,46; 0,43 e 0,37, respectivamente), que decresceram até próximo à 6ª semana (0,34, aproximadamente). Desse ponto seguiram a mesma tendência das estimadas pelos modelos A4C6R4 e A4C5R4: aumentaram até próximo ao pico (variando de 0,27 na primeira semana a 0,37, próximo ao pico). Do pico à 18ª semana, as estimativas foram semelhantes. Desse ponto até o final da lactação, as herdabilidades estimadas tenderam a decrescer, enquanto as estimadas pelos modelos A4C7R4 e A5C7R4 tiveram comportamento ligeiramente diferenciado. O modelo A5C7R4 estimou herdabilidades decrescentes, como os outros, porém mais altas até a 38ª semana (0,24). O modelo A4C7R4, por sua vez, estimou herdabilidades crescentes a partir da 36ª e chegou a estimar a herdabilidade mais alta da última semana (0,25), quando as herdabilidades estimadas pelos modelos estudados variaram de 0,12 a 0,25.

Tanto os modelos A4C5R4 e A4C6R4 quanto o modelo animal tradicional estimaram herdabilidades mais altas no meio da lactação. Esse mesmo comportamento tem sido relatado em estudos com bovinos leiteiros utilizando os modelos ordinários para produção de leite no dia do controle (Ferreira et al., 2003; Mello et al., 2005), o que sugere maior

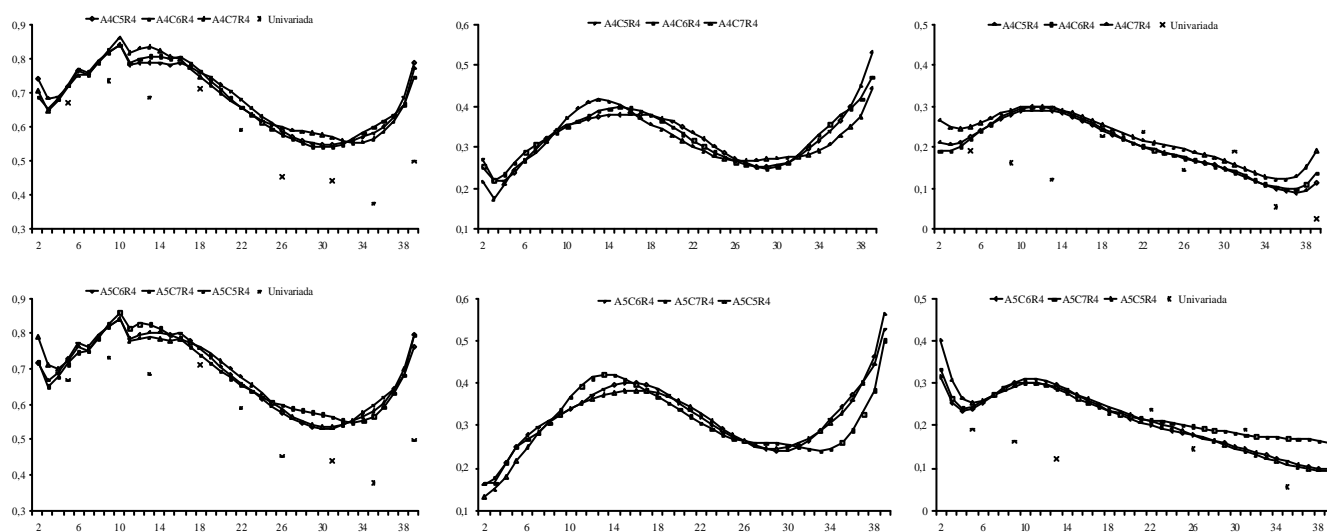


Figura 2 - Estimativas das variâncias fenotípica (à esquerda), de ambiente permanente (no meio) e genética aditiva (à direita) para as produções de leite semanais obtidas por modelos selecionados.

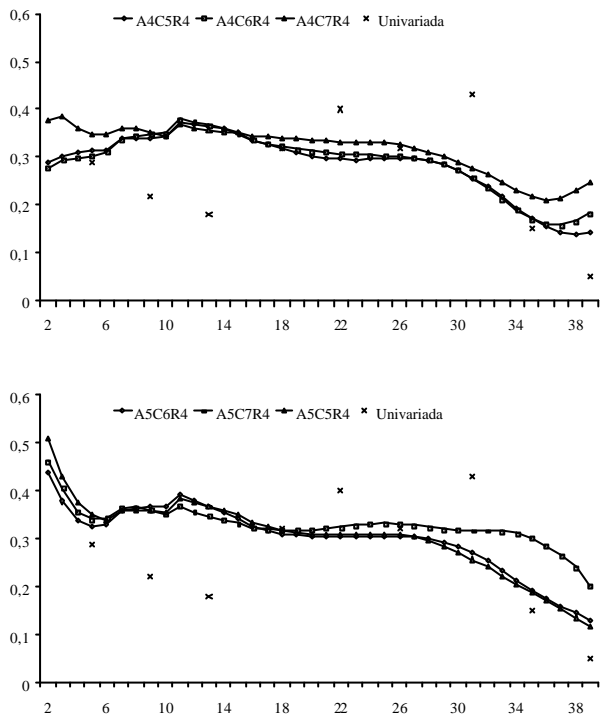


Figura 3 - Estimativas de herdabilidade para as produções de leite semanais obtidas por modelos selecionados.

susceptibilidade às variações ambientais nos controles do início e final da lactação e que as produções do meio da lactação são mais influenciadas pelas diferenças genéticas e permanentes de meio existentes entre as vacas, como observado nas cabras. De modo geral, as herdabilidades estimadas para as produções de leite semanais foram maiores que as encontradas na literatura nacional para produção total na lactação (Tholon et al., 2001; Gonçalves et al., 2002; Pimenta Filho et al., 2004). Na literatura internacional, em poucos trabalhos constam estimativas de herdabilidade para produções no dia do controle em cabras, como o trabalho de Breznik et al. (2000), que estimaram herdabilidade por meio de um modelo de repetibilidade igual a 0,20.

As estimativas de correlações genéticas, com exceção das estimadas pelo modelo A5C7R4, foram todas positivas, embora algumas diferenças tenham ocorrido entre as estimativas dos demais modelos selecionados. As estimadas pelos modelos A4C5R4 e A5C5R4 descreveram uma superfície ondulosa, com alguns pontos em que as correlações não diminuíram com o aumento das semanas entre os controles, embora o modelo com ordem 5 para o efeito genético aditivo tendeu a estimar uma superfície mais

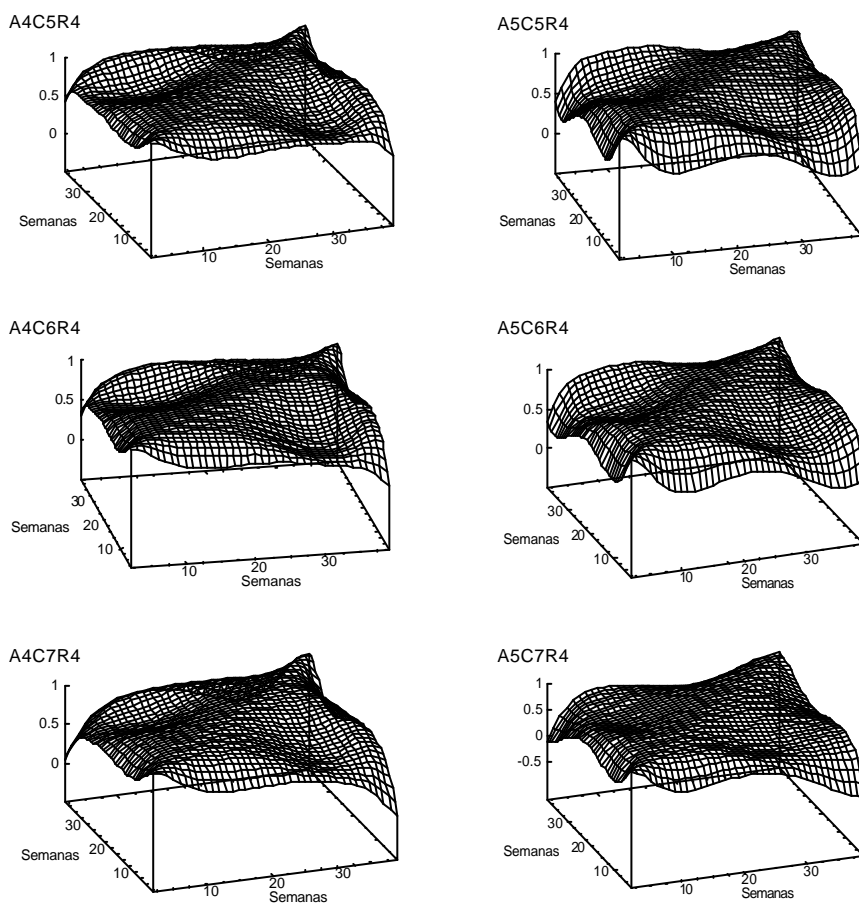


Figura 4 - Correlações genéticas estimadas entre as produções de leite semanais por modelos selecionados.

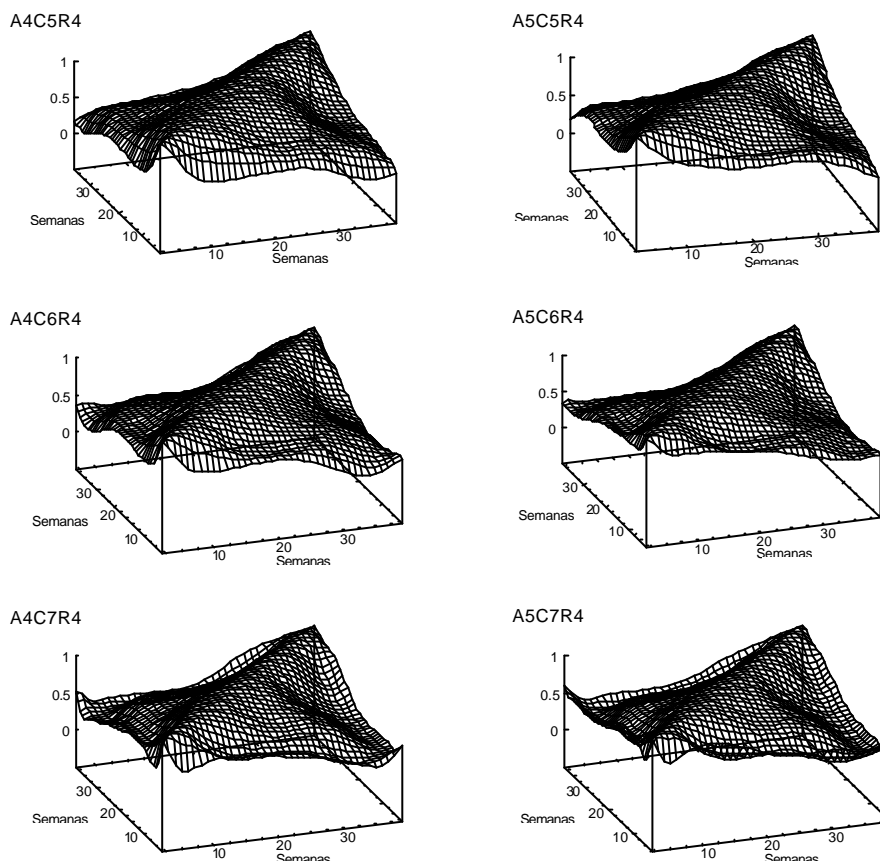


Figura 5 - Correlações de ambiente permanente estimadas entre as produções de leite semanais por modelos selecionados.

ondulada nas extremidades (Figura 4). Quando a ordem de ajuste para o efeito de ambiente permanente aumentou de 5 (A4C5R4 e A5C5R4) para 6 (A4C6R4 e A5C6R4), as superfícies de correlações genéticas foram equivalentes. Entretanto, ao ajustar uma função de ordem 7 (A4C7R4 e A5C7R4), observaram-se mudanças mais expressivas no padrão de correlações genéticas entre os controles semanais. A superfície estimada pelo modelo A4C7R4 descreveu correlações genéticas próximas de zero entre a primeira e a última semana, ao passo que o modelo A5C7R4 chegou a estimar correlações negativas entre controles no início e final da lactação, o que não é condizente com o fenômeno biológico em estudo. Além disso, esse resultado comprova que a modelagem do efeito de ambiente permanente pode interferir na modelagem do genético aditivo, dependendo da ordem de ajuste do polinômio.

De modo geral, as correlações genéticas estimadas foram de magnitudes altas e indicam que seleção em qualquer ponto da curva pode promover resposta correlacionada em outros pontos e na mesma direção.

A tendência das correlações para o efeito de ambiente permanente para a ordem mais elevada, sétima, não atendeu

à expectativa, ou seja, as correlações apresentaram superfície ondulada com tendência de aumento nas das correlações nas extremidades da curva (Figura 5). As correlações estimadas pelos modelos com ordem 5 e 6 foram mais suaves.

Assim, de acordo com os critérios Log L, AIC e BIC, os modelos A5C7R4 e A4C7R4 proporcionaram melhor ajuste para as variâncias genética e de ambiente permanente, todavia, o modelo A4C6R4, ou um menos parametrizado, A4C5R4, estimou herdabilidades e correlações genéticas mais próximas do biologicamente esperado, sugerindo que foram suficientes para estimação de parâmetros genéticos para produção de leite no dia do controle em caprinos. Além disso, pelo comportamento das variâncias estimadas e dos autovalores, é possível questionar a utilização dos critérios adotados inicialmente como as únicas ferramentas utilizadas para avaliar qualidade de ajuste em modelos de regressão aleatória.

Ressalta-se também que os resultados obtidos neste estudo devem ser analisados com cautela, em virtude do pequeno número de observações, o que torna necessária a realização de nova pesquisa com amostras mais representativas.



## Conclusões

A variância residual homogênea ao longo da curva de lactação é inadequada e deve ser considerada heterogênea com pelo menos quatro classes de variâncias heterogêneas. O modelo de quinta ordem para a variância genética aditiva e sétima para o ambiente permanente é indicado para modelar as variâncias genéticas e de ambiente permanente. Entretanto, um modelo mais parcimonioso, de quarta ordem para o efeito genético aditivo e de sexta ordem para o ambiente permanente, é suficiente para modelar as mudanças nas variâncias ao longo da curva de lactação de cabras.

## Agradecimento

Ao doutorando em Genética e Melhoramento Animal da UFV Marcos Yamaki, pela valiosa contribuição na realização deste trabalho.

## Literatura Citada

- ARAÚJO, C.V.; TORRES, R.A.; COSTA, C.N. et al. Uso de funções ortogonais para descrever a produção de leite no dia de controle por meio de modelos de regressão aleatória. **Revista Brasileira de Zootecnia**, v.35, n.3, p.967-974, 2006.
- ASSIS, G.M.L.; ALBUQUERQUE, L.G.; SARMENTO, J.L.R. et al. Estimação de parâmetros genéticos em caprinos leiteiros por meio da análise de regressão aleatória, utilizando a amostragem de gibbs. **Revista Brasileira de Zootecnia**, v.35, n.3, p.706-714, 2006.
- BREZNIK, S.; MALOVRH, S.; KOVAC, M. et al. Additive genetic and environmental variance components for milk traits in goat with test day model. **Zootecnika**, v.76, n.1, p.61-66, 2000.
- DRUET, T.; JAFFREZIC, F.; BOICHARD, D. et al. Modeling lactation curves and estimation of genetic parameters for first lactation test-day records of French Holstein cows. **Journal Dairy Science**, v.86, p.2480-2490, 2003.
- EL FARO, L.; ALBUQUERQUE, L.G. Utilização de modelos de regressão aleatória para a produção de leite no dia do controle, com diferentes estruturas de variâncias residuais. **Revista Brasileira de Zootecnia**, v.32, n.5, p.1104-1113, 2003.
- FERREIRA, W.J.; TEIXEIRA, N.M.; EUCLYDES, R.F. et al. Avaliação genética de bovinos da raça Holandesa usando a produção de leite no dia do controle. **Revista Brasileira de Zootecnia**, v.32, n.2, p.295-303, 2003.
- FOULLEY, J.L.; ROBERT-GRANIE, C. Heteroskedastic random coefficient models. In: WORLD CONGRESS OF GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier. **Proceedings...** Montpellier: 2002. (CD-ROM).
- GONÇALVES, H.C.; SILVA, M.A.; WECHSLER, F.S. et al. Parâmetros e tendência genética da produção de leite de cabra no Brasil. **Revista Brasileira de Zootecnia**, v.31, n.6, p.2204-2208, 2002.
- HENDERSON JR., C.R. Analysis of covariance in the mixed model: higher-level, nonhomogeneous, and random regressions. **Biometrics**, v.38, p.623-640, 1982.
- KETTUNEN, A.; MANTYSAARI, E.A.; STRANDEN, I. et al. Estimation of genetic parameters for first lactation test day milk production using random regression model. In: WORLD CONGRESS OF GENETICS APPLIED TO LIVESTOCK PRODUCTION, 6., 1998, Armidale. **Proceedings...** Armidale: University of New England, 1998. (CD-ROM).
- KIRKPATRICK, M.; LOFSVOLD, D.; BULMER, M. Analysis of the inheritance, selection and evolution of growth trajectories. **Genetics**, v.124, p.979-993, 1990.
- LEGARRA, A.; MISZTAI, I.; BERTRAND, J.K. Constructing covariance functions for random regression models for growth in Gelbvieh beef cattle. **Journal Animal Science**, v.82, p.1564-1571, 2004.
- LÓPEZ-ROMERO, P.; CARABAÑO, M.J. Comparing alternative random regression models to analyse first-lactation daily milk yield data in Holstein-Friesian cattle. **Livestock Production Science**, v.82, p.81-86, 2003.
- MELLO, C.M.R.; PACKER, I.U.; COSTA, C.N. et al. Parâmetros genéticos para as produções de leite no dia do controle e da primeira lactação de vacas da raça Holandesa. **Revista Brasileira de Zootecnia**, v.34, n.3, p.796-806, 2005.
- MEYER, K. DXMRR – A program to estimate covariance functions for longitudinal data by REML. In: WORLD CONGRESS OF GENETICS APPLIED TO LIVESTOCK PRODUCTION, 6., 1998, Armidale. **Proceedings...** Armidale: University of New England, 1998. (CD-ROM).
- ØDEGARD J.; JENSEN, J.; KLEMETS DAL, G. et al. Genetic analysis of somatic cell score in Norwegian cattle using random regression test-day models. **Journal Dairy Science**, v.86, p.4103-114, 2003.
- PIMENTA FILHO, E.C.; SARMENTO, J.L.R.; RIBEIRO, M.N. Efeitos genéticos e ambientais que afetam a produção de leite e duração da lactação de cabras mestiças no estado da Paraíba. **Revista Brasileira de Zootecnia**, v.31, n.6, p.1426-1431, 2004.
- RE KAYA, R.; CARABAÑO, M.J.; TORO, M.A. Assessment of heterogeneity of residual variance using changing points techniques. **Genetics Selection Evolution**, v.32, p.339-346, 2000.
- SARMENTO, J.L.R.; REIS FILHO, J.C.; ALBUQUERQUE, L.G. et al. Avaliação genética de caprinos usando a produção de leite no dia de controle. **Revista Brasileira de Zootecnia**, v.35, n.2, p.443-451, 2006.
- SCHAEFFER, L.R. Application of regression models models in animal breeding. **Livestock Production Science**, v.86, p.35-45, 2004.
- STRABEL, T.; MISZTAI, I. Genetic parameters for first and second lactation milk yields of Polish Black and White cattle with random regression test-day models. **Journal Dairy Science**, v.82, p.2805-2810, 1999.
- SULLIVAN, B.; WIGGANS, G. Genetic evaluation of dairy goats in the United States and Canada. In: INTERNATIONAL CONFERENCE ON GOATS, 7., 2000, Poitiers. **Proceedings...** Poitiers: 2000. (CD-ROM).
- THOLON, P.; QUEIROZ, S.A.; RIBEIRO, A.C. et al. Estudo genético quantitativo da produção de leite em caprinos da raça Saanen. **Archivos Latinoamericanos Producción Animal**, v.9, p.1-5, 2001.
- Van Der WERF, J.H.J.; GODDARD, M.E.; MEYER, K. The use of covariance functions and random regressions for genetic evaluation of milk production based on test day records. **Journal Dairy Science**, v.81, p.3300-3308, 1998.