



Estimação de parâmetros genéticos em caprinos leiteiros por meio de análise de regressão aleatória utilizando-se a Amostragem de Gibbs

Giselle Mariano Lessa de Assis¹, Lucia Galvão de Albuquerque², José Lindenberg Rocha Sarmiento³, José Marques Carneiro Júnior⁴, Paulo Sávio Lopes⁵, Marcelo Teixeira Rodrigues⁵

¹ Embrapa Acre, Rodovia BR 364, km 14, CP 321, CEP: 69908-970, Rio Branco - AC.

² FCAV/UNESP/Departamento de Zootecnia, CEP: 14884-900, Jaboticabal, SP.

³ Doutorando em Genética e Melhoramento da UFV, CEP: 36571-000, Viçosa, MG.

⁴ Bolsista DCR do CNPq/FUNTAC da Embrapa Acre, Rodovia BR 364, km 14, CEP: 69908-770, Rio Branco, AC.

⁵ Departamento de Zootecnia/UFV, CEP: 36571-000, Viçosa, MG.

RESUMO - Modelos de regressão aleatória foram utilizados neste estudo para estimar parâmetros genéticos da produção de leite no dia do controle (PLDC) em caprinos leiteiros da raça Alpina, por meio da metodologia Bayesiana. As estimativas geradas foram comparadas às obtidas com análise de regressão aleatória, utilizando-se o REML. As herdabilidades encontradas pela análise Bayesiana variaram de 0,18 a 0,37, enquanto, pelo REML, variaram de 0,09 a 0,32. As correlações genéticas entre dias de controle próximos se aproximaram da unidade, decrescendo gradualmente conforme a distância entre os dias de controle aumentou. Os resultados obtidos indicam que: a estrutura de covariâncias da PLDC em caprinos ao longo da lactação pode ser modelada adequadamente por meio da regressão aleatória; a predição de ganhos genéticos e a seleção de animais geneticamente superiores é viável ao longo de toda a trajetória da lactação; os resultados gerados pelas análises de regressão aleatória utilizando-se a Amostragem de Gibbs e o REML foram semelhantes, embora as estimativas das variâncias genéticas e das herdabilidades tenham sido levemente superiores na análise Bayesiana, utilizando-se a Amostragem de Gibbs.

Palavras-chave: análise Bayesiana, caprinocultura leiteira, componentes de variância, correlação genética

Estimation of genetic parameters for milk yield of dairy goats by random regression analysis using Gibbs Sampling

ABSTRACT - Random regression models were used to estimate genetic parameters for test-day milk yield (PLDC) of Alpine dairy goats, implemented by Bayesian methods with Gibbs Sampling. The estimates were compared with those obtained by random regression analysis, using REML. Heritability estimates obtained by Bayesian analysis ranged from 0.18 to 0.37, while those obtained by REML ranged from 0.09 to 0.32. Genetic correlations between yields of close test days approached the unit, but decreased gradually as the interval between test days increased. Results indicated that random regression models are appropriate to model the covariance structure of PLDC and to predict genetic gains and select animals along the lactation trajectory of dairy goats. Results obtained by Bayesian and REML approaches were similar, although genetic variance and heritability estimates were slightly higher with Bayesian methods.

Key Words: Bayesian methods, dairy goats, genetic correlation, variance components

Introdução

A caprinocultura leiteira é uma atividade potencialmente rentável, que tem desenvolvido significativamente nas últimas três décadas no Brasil. O rebanho caprino brasileiro possui aproximadamente 12,8 milhões de cabeças, constituindo-se o nono maior rebanho do mundo, porém, contribuindo com apenas 1,3% da produção de leite de cabra (FAO, 2000). O grande pólo de produção de caprinos localiza-se na Região Nordeste, que engloba 90% do rebanho nacional. Na Região Sudeste, por sua vez, há em torno de 3% do rebanho, formado

principalmente pelas raças Saanen, Alpina e Toggenburg, selecionadas para produção de leite (Souza et al., 2002).

Embora essas raças estejam presentes há algumas décadas no país, pouco se conhece sobre seu desempenho e são escassas as estimativas de parâmetros genéticos das características produtivas que apresentam número consistente de observações. A melhoria dos plantéis se baseia na importação de animais e de sêmen, principalmente da França, onde se realiza a avaliação genética de reprodutores (Gonçalves et al., 2002). Porém, a seleção praticada em outro país/região pode não atender às necessidades dos produ-

tores brasileiros. Assim, a geração de conhecimento para viabilizar o melhoramento genético de caprinos no Brasil é uma necessidade e um desafio constante às instituições de pesquisa.

No melhoramento genético, freqüentemente se selecionam animais com base em características medidas no mesmo indivíduo, em diferentes estádios fisiológicos, o que merece atenção especial, visto que diferentes genes podem contribuir para expressão da característica nesses diferentes estágios. Na avaliação quantitativa de medidas repetidas, quatro classes de modelos de análises podem ser empregadas: modelos univariados para cada medida no tempo (dias em lactação, idade, entre outras), modelos de repetibilidade, modelos multicaracterísticos e modelos de regressão aleatória. Os modelos usualmente empregados (univariado, de repetibilidade e multicaracterístico) têm sido questionados quanto à sua praticidade e adequabilidade às situações encontradas. Modelos alternativos de análise, ou seja, que corrijam as limitações daqueles em uso, têm sido propostos, entre eles, o de regressão aleatória (Jamrozik & Schaeffer, 1997).

Quando se ajusta um modelo de regressão aleatória, assume-se, implicitamente, certa estrutura de covariância entre as observações, determinada pelas covariâncias entre os coeficientes de regressão e caracterizada como uma função de covariância (Kirkpatrick et al., 1990). A função de covariância, uma função bivariada, é a análoga da matriz de covariâncias, amplamente empregada na estatística multivariada e na genética quantitativa.

O uso de funções de covariância é vantajoso porque cria uma estrutura de covariância mais acurada das características em estudo, possibilita a predição de estruturas de covariância para qualquer ponto ao longo de uma escala contínua, permite a utilização de medidas tomadas em qualquer momento da trajetória, sem a necessidade de se fazer correções para idades, e fornece uma metodologia para analisar padrões de covariância associados a mudanças particulares da característica ao longo da trajetória.

Existem diversas opções para se estimarem os coeficientes de funções de covariância, destacando-se o Método da Máxima Verossimilhança Restrita (REML) ou os métodos Bayesianos, com base em Cadeia de Markov Monte Carlo (MCMC). No REML, normalmente, são empregados os seguintes algoritmos: DF (*derivative-free*), M (*expectation-maximization*) ou AI (*average information*). Para modelos de grandes dimensões, os algoritmos DF e EM podem requerer muitos ciclos de iterações para atingir a convergência, enquanto o algoritmo AI usualmente requer menos que dez ciclos (Madsen, 1994). Entretanto, a demanda

computacional por ciclo para os três algoritmos é de magnitude similar, ocorrendo pequeno acréscimo dessa demanda quando mais derivadas são usadas. A cada iteração, a maior demanda computacional nos três algoritmos REML relaciona-se à criação e fatoração das equações de modelos mistos de Henderson (Jensen, 2001).

Com métodos MCMC, que usam Amostragem de Gibbs, o problema de criar e fatorar as equações de modelos mistos pode ser contornado (Geman & Geman, 1984). Além disso, os métodos baseados na estatística Bayesiana possuem a vantagem de permitir a inclusão na análise do conhecimento *a priori* existente sobre os parâmetros desconhecidos (Gianola & Fernando, 1986). A demanda computacional de um ciclo de iteração desses métodos é menor que em um ciclo dos métodos baseados no REML. Requerimentos de memória para o algoritmo são também de magnitude similar, possibilitando analisar maior conjunto de dados.

A análise Bayesiana utilizando-se a Amostragem de Gibbs, portanto, consiste em uma alternativa mais simples para ser implementada que o REML e, como não é limitação o fato de necessitar de maior número de ciclos para obtenção das amostras *a posteriori*, tem sido recomendada para estimação de parâmetros genéticos da produção de leite no dia do controle em gado de leite (Misztal et al., 2000).

Neste trabalho, utilizou-se metodologia Bayesiana com os objetivos de estimar, em caprinos leiteiros, parâmetros genéticos da produção de leite no dia do controle (PLDC) por meio de regressão aleatória e comparar essas estimativas às obtidas na análise de regressão aleatória, por meio do REML.

Material e Métodos

Foram analisados 9.374 registros da produção de leite no dia do controle (PLDC), provenientes da primeira lactação de 302 cabras da raça Alpina, no período de 1997 a 2004, pertencentes ao rebanho do Departamento de Zootecnia da Universidade Federal de Viçosa (Viçosa – MG). A partir de um banco de dados composto de 17.482 registros, foram excluídos das análises animais com mais de 36 meses de idade, com três desvios-padrão acima ou abaixo da média de produção de leite da semana de lactação ou com menos de seis controles durante a lactação considerada. A média do número de animais dentro dos 138 grupos contemporâneos formados (definido pelas variáveis ano e quinzena do controle) foi igual a 38, sendo o número mínimo igual a cinco e o máximo igual a 78 animais.

Os animais foram mantidos em baias coletivas sob sistema de estabulação livre e receberam alimentação à base de silagem de milho e feno, além de mistura concentrada,

fornecida conforme as necessidades dos animais. As cabras foram submetidas à ordenha mecânica duas vezes ao dia. A PLDC foi considerada em classes semanais, sendo utilizados nas análises os controles obtidos entre a 2ª e a 39ª semanas de lactação. O número de registros, assim como a média e o desvio-padrão para produção de leite referente à semana do controle, encontram-se na Tabela 1.

As estimativas dos componentes de (co)variâncias foram obtidas pela análise Bayesiana (AG-RA), utilizando-se o programa computacional RRGIBBS (Meyer, 2002). Este programa gera Cadeias de Markov para os parâmetros de um modelo de regressão aleatória, por meio da Amostragem de Gibbs. Uma única cadeia composta por 200.000 amostras foi gerada para cada análise, descartando-se as 10.000 primeiras amostras correspondentes ao período de *burn-in*. *Flat priors* foram designados para a matriz de (co)variâncias dos efeitos aleatórios genético-aditivos e de ambiente permanente e para os componentes de variâncias residuais.

A análise visual das amostras geradas a cada iteração para cada parâmetro desconhecido e o programa GIBANAL (Van Kaam, 1998) foram utilizados para definir o tamanho da cadeia, o período de *burn-in* e o intervalo de amostragem (*thinning interval*) para obtenção das densidades marginais *a posteriori*.

As estimativas dos componentes de (co)variância foram obtidas com base na média *a posteriori*, considerando-se 190.000 amostras. Foram calculados percentis de interesse para construção de intervalos de credibilidade, ou seja, conjuntos de valores que contêm o verdadeiro valor

paramétrico com alta probabilidade (Sorensen & Gianola, 2002). O intervalo de credibilidade construído foi de tamanho $1 - \alpha$, sendo α igual a 0,05. Assim, foi determinada a região da distribuição *a posteriori* de alta probabilidade (*highest posterior density*), compreendida entre os percentis 2,5 e 97,5.

As análises de regressão aleatória foram conduzidas considerando-se um modelo animal unicaracterístico, aplicado a registros da PLDC, como a seguir:

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

em que y_{ij} = produção de leite no j éssimo controle da i éssima cabra; S = conjunto de efeitos sistemáticos, constituído pelo efeito de *grupo contemporâneo* (definido por ano/quinzena do controle), composto por 138 subclasses e pelos efeitos linear e quadrático da covariável *idade da cabra*; $b_m = m$ éssimo coeficiente de regressão da PLDC sobre a semana de lactação para a média da população; a_{im} e $g_{im} = m$ éssimos coeficientes de regressão genético-aditivo e de ambiente permanente, respectivamente, para a i éssima cabra; k_b , k_a e k_g = ordens dos polinômios a serem ajustados; $(t_{ij}) = j$ éssima semana de lactação da i éssima cabra; $f_m(t_{ij}) = m$ éssima função polinomial a ser avaliada para (t_{ij}) ; e_{ij} = efeito de ambiente temporário afetando y_{ij} .

Na forma matricial, o modelo [1] pode ser escrito como:

$$y = Xb + Za + Wc + e, \quad [2]$$

em que b = vetor que contém os efeitos sistemáticos e os

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

Table 1 - Number of records (NR), mean and standard deviation (DP) of milk yield at the week of test day in first lactation of Alpine goats

Semana Week	NR	Média (kg) Mean (kg)	DP	Semana Week	NR	Média (kg) Mean (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

coeficientes b_m da regressão fixa; \mathbf{a} = vetor contendo os coeficientes \mathbf{a}_m da regressão aleatória, que descreve a trajetória do valor genético aditivo para cada animal; \mathbf{c} = vetor contendo os coeficientes \mathbf{g}_m da regressão aleatória que descreve a trajetória dos efeitos de ambiente permanente para cada animal; \mathbf{e} = vetor que contém os efeitos residuais; \mathbf{X} , \mathbf{Z} e \mathbf{W} = matrizes de incidência e de covariáveis.

Pressupõe-se que:

$$\bullet \mathbf{y} | \mathbf{b}, \mathbf{a}, \mathbf{c}, \mathbf{S}_{el}^2, \dots, \mathbf{S}_{ef}^2 \sim NMV(\mathbf{Xb} + \mathbf{Za} + \mathbf{Wc}, \mathbf{R}), \text{ com}$$

$\mathbf{R} = \text{diag} \{ \mathbf{S}_{el}^2 \}$, sendo essa a função de verossimilhança.

As pressuposições acerca das distribuições *a priori* para os parâmetros desconhecidos são consideradas como a seguir (Rekaya et al., 1999):

• $\mathbf{a} | \mathbf{S}_a \sim NMV(\mathbf{0}, \mathbf{G})$, com $\mathbf{G} = \mathbf{A} \otimes \mathbf{S}_a$, em que \mathbf{A} é a matriz de parentesco entre os animais e \mathbf{S}_a , a matriz contendo as (co)variâncias entre os coeficientes de regressão aleatória genético-aditivos;

• $\mathbf{c} | \mathbf{S}_c \sim NMV(\mathbf{0}, \mathbf{C})$, com $\mathbf{C} = \mathbf{I} \otimes \mathbf{S}_c$, em que \mathbf{I} é uma matriz identidade e \mathbf{S}_c é a matriz contendo as (co)variâncias entre os coeficientes de regressão aleatória de ambiente permanente;

$$\bullet \mathbf{S}_a | n_a, \mathbf{S}_a^2 \sim \mathbf{W}^{-1}(n_a, n_a \mathbf{S}_a^2);$$

• $\mathbf{S}_c | n_c, \mathbf{S}_c^2 \sim \mathbf{W}^{-1}(n_c, n_c \mathbf{S}_c^2)$, em que n_a, \mathbf{S}_a^2 e n_c, \mathbf{S}_c^2 representam o grau de confiança e os valores *a priori* para as covariâncias dos coeficientes de regressão genético-aditivos e de ambiente permanente, respectivamente;

• $\mathbf{S}_{el}^2 | n_{el}, \mathbf{S}_{el}^2 \sim \mathbf{C}^{-2}(n_{el}, n_{el} \mathbf{S}_{el}^2)$, com $l = 1, 2, \dots, f$, em que f é o número de classes de resíduo e

$$\bullet \mathbf{b} \propto \text{constante}.$$

NMV , \mathbf{W}^{-1} e \mathbf{C}^{-2} referem-se às distribuições normal multivariada, Wishart invertida e Qui-quadrado invertida escalonada, respectivamente. Detalhes sobre as distribuições condicionais completas *a posteriori* foram descritos por Jamrozik & Schaeffer (1997) e Rekaya et al. (1999).

Foi considerada uma regressão sobre o tempo (semana de lactação), o que gerou a trajetória média da população (fixa), modelada por polinômios ortogonais de Legendre, com ordem de ajuste igual a 3 para todos os modelos analisados. Foram também ajustados polinômios ortogonais de Legendre para a regressão dos efeitos aleatórios genético-aditivo e de ambiente permanente. Diferentes ordens de ajuste para a regressão da parte aleatória do modelo foram consideradas e são apresentadas na Tabela 2, juntamente com o número de classes do resíduo, considerado heterogêneo para cada modelo estudado. A variância residual pôde variar

Tabela 2 - Ordens de ajuste dos polinômios de Legendre para as regressões aleatórias genético-aditiva (k_a) e de ambiente permanente (k_g) e o número de classes do resíduo (f) para cada modelo considerado

Table 2 - Orders of fit of Legendre polynomials for additive genetic (k_a) and permanent environmental (k_g) effects and number of residual classes (f) for each random regression model

Modelo Model	k_a	k_g	Classe do resíduo Residual class
PL ₃₃₂	3	3	2
PL ₃₃₄	3	3	4
PL ₃₅₄	3	5	4
PL ₄₄₄	4	4	4
PL ₄₅₄	4	5	4

entre grupos de semanas, porém, dentro de cada grupo, a variância residual foi considerada a mesma. Foram formadas duas classes de resíduo (constituídas pelas semanas 2 a 15 e 16 a 39) e quatro classes de resíduo (constituídas pelas semanas 2 a 6, 7 a 10, 11 a 15 e 16 a 39).

Os diferentes modelos utilizados foram comparados por meio das estimativas da variância residual (Jamrozik & Schaeffer, 2002), sendo apresentados os resultados daquele que melhor se ajustou. Esse método utiliza as médias ponderadas (pelo número de semanas em cada classe) das estimativas das variâncias residuais, conforme a expressão:

$$VRP = \frac{1}{TS} \sum_{l=1}^f \hat{\mathbf{S}}_{el}^2 NS_l, \quad [3]$$

em que $\hat{\mathbf{S}}_{el}^2$ = variância residual estimada para a *l*ésima classe; f = número de classes de resíduos para cada modelo; NS_l = número de semanas existente na *l*ésima classe; e TS = número total de semanas em lactação considerado no estudo.

Realizaram-se análises semelhantes às descritas, utilizando-se, porém, para estimação dos componentes de (co)variâncias, o Método da Máxima Verossimilhança Restrita (REML-RA), por meio da opção DXMRR do programa DFREML (Meyer, 1998). Considerou-se na condução dessas análises o mesmo modelo descrito em [1] e [2], fazendo-se as seguintes pressuposições:

$$\mathbf{E} \begin{bmatrix} \mathbf{y} \\ \mathbf{a} \\ \mathbf{c} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{Xb} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \text{ em que } \text{Var}(\mathbf{a}) = \mathbf{G}, \text{Var}(\mathbf{c}) = \mathbf{C} \text{ e } \text{Var}(\mathbf{e}) = \mathbf{R},$$

conforme definido anteriormente.

Com o intuito de verificar a tendência das curvas apresentadas na Figura 1 em relação às estimativas obtidas na análise que desconsidera a estrutura de correlação existente entre as

medidas repetidas ao longo da vida do mesmo indivíduo, realizaram-se análises Bayesianas unicaracterísticas (AG-UNI), utilizando-se o programa MTGSAM – *Multiple Trait Gibbs Sampling Analyses in Animal Models* (Van Tassell & Van Vleck, 1995), que também gera Cadeias de Markov Monte Carlo, por meio da Amostragem de Gibbs.

Foram realizadas nove análises, nas quais se considerou, como diferentes características, a PLDC medida em distintos dias da lactação. Assim, as nove características analisadas foram a PLDC aos 30, 60, 90, 120, 150, 180, 210, 240 e 270 dias. Foram geradas cadeias com 300.000 amostras e o período de *burn-in* referente a 30.000 iterações.

Matricialmente, o modelo utilizado foi o seguinte:

$$y = Xb + Za + e,$$

em que y = vetor contendo as observações da PLDC; b = vetor que contém os efeitos *fixos*, contendo grupo contemporâneo e a covariável idade da cabra; a = vetor de efeitos aleatórios genético-aditivos; e = vetor de resíduos aleatórios; X e Z = as matrizes de incidência que relacionam as observações aos efeitos *fixos* e aleatórios genético-aditivos, respectivamente. As pressuposições assumidas em relação aos vetores a e e são as que possuem distribuição normal, com $E(a) = E(e) = 0$ e $Var(a) = A \otimes S_a^2$ e $Var(e) = I \otimes S_e^2$, sendo S_a^2 e S_e^2 as variâncias genético-aditiva e residual, respectivamente; A , a matriz de numeradores dos coeficientes de parentesco de Wright entre os animais; e I , uma matriz identidade.

Resultados e Discussão

As Variâncias Residuais Ponderadas (VRP) para os modelos PL₃₃₂, PL₃₃₄, PL₃₅₄, PL₄₄₄ e PL₄₅₄ foram, respectivamente, 0,17; 0,18; 0,16; 0,17 e 0,16, indicando como de melhor ajuste os modelos PL₃₅₄ e PL₄₅₄.

O logaritmo da função de máxima verossimilhança (LMV), o Critério de Informação de Akaike (AIC) e o Critério de Informação Bayesiano de Schwarz (BIC), largamente empregados para comparação de modelos em análises frequentistas, encontram-se na Tabela 3. Menores valores para AIC e BIC indicam melhor ajuste, de modo que ambos os testes permitem comparação entre modelos não aninhados e penalizam aqueles com maior número de parâmetros. Os modelos indicados como de melhor ajuste pela análise REML-RA, conforme os critérios da Tabela 3, foram os mesmos indicados pela VRP (PL₃₅₄ e PL₄₅₄). Uma vez que o modelo PL₃₅₄ é o menos parametrizado, optou-se por apresentar seus resultados.

Tabela 3 - Log da função de máxima verossimilhança (LMV), Critério de Informação de Akaike (AIC) e Critério de Informação Bayesiano (BIC) para os diferentes modelos

Table 3 - Values of log likelihood function (LMV), Akaike Criterion Information (AIC) and Bayesian Criterion Information (BIC) for different models

Modelo Model	LMV	AIC	BIC
PL ₃₃₂	2055	-4081	-3981
PL ₃₃₄	2150	-4267	-4152
PL ₃₅₄	2386	-4722	-4543
PL ₄₄₄	2332	-4616	-4444
PL ₄₅₄	2405	-4751	-4544

As estimativas das médias, os desvios-padrão e os percentis da distribuição *a posteriori* das (co)variâncias genéticas e de ambiente permanente entre os coeficientes de regressão aleatória, assim como as variâncias residuais, considerando o modelo PL₃₅₄, encontram-se na Tabela 4. Os componentes de variância para os primeiros coeficientes genético-aditivos e de ambiente permanente da regressão aleatória foram os que apresentaram maior magnitude (0,3967 e 0,4807, respectivamente). Os desvios-padrão das covariâncias foram, em geral, elevados, como consequência do número relativamente reduzido de animais na análise. O desvio da média *a posteriori* em relação à respectiva mediana da distribuição (percentile 50,0%), fornece informação sobre a simetria da marginal *a posteriori*. De modo geral, para todos os parâmetros estimados, os valores foram próximos. Os maiores desvios, porém, ocorreram para as (co)variâncias dos coeficientes genético-aditivos. Na última coluna da Tabela 4 são apresentadas as estimativas pontuais obtidas pela metodologia REML-RA. As estimativas geradas para todos os parâmetros foram próximas às obtidas pelo AG-RA e encontram-se todas no intervalo Bayesiano de credibilidade.

As estimativas das variâncias fenotípicas, genéticas, de ambiente permanente e residuais, assim como das herdabilidades e das frações da variância fenotípica referente à variância de ambiente permanente (c^2) e da PLDC obtida para o modelo PL₃₅₄, encontram-se na Figura 1. A variância total da PLDC é a soma das variâncias genética, de ambiente permanente e residual, de modo que cada um desses componentes pode apresentar heterogeneidade de variância ao longo da curva de lactação, como constatado em diversos estudos com vacas de leite (Olori et al., 1999; Brotherstone et al., 2000; Rekaya et al., 2000).

As variâncias fenotípicas variaram de 0,58 a 0,95 kg² quando obtidas pelo AG-RA e de 0,54 a 0,85 kg² quando estimadas pelo REML-RA. O comportamento de sua trajetória refletiu os resultados encontrados para cada um de seus

Tabela 4 - Média, desvio-padrão (DP) e percentis das distribuições *a posteriori* das (co)variâncias dos coeficientes de regressão aleatória genético-aditiva (\mathbf{a}_i \mathbf{a}_j) e de ambiente permanente (\mathbf{g}_i \mathbf{g}_j) e das variâncias residuais (e_i e_j), pelo o modelo PL₃₅₄. Na última coluna são apresentadas as estimativas pontuais obtidas pela metodologia REML-RA

Table 4 - Mean, standard-deviation (DP) and percentiles of random regression (co)variance coefficients of genetic-additive (\mathbf{a}_i \mathbf{a}_j) and permanent environmental (\mathbf{g}_i \mathbf{g}_j) effects, residual variances (e_i e_j) *a posteriori* distributions for model PL₃₅₄ and respective estimates obtained by REML-RA

Parâmetro <i>Parameter</i>	Média <i>Mean</i>	DP	2,5%	50,0%	97,5%	REML - RA
\mathbf{a}_1 \mathbf{a}_1	0,3967	0,150	0,1558	0,3709	0,7616	0,3308
\mathbf{a}_1 \mathbf{a}_2	-0,0331	0,046	-0,1304	-0,0301	0,0499	-0,0357
\mathbf{a}_1 \mathbf{a}_3	-0,0350	0,028	-0,0960	-0,0327	0,0143	-0,0336
\mathbf{a}_2 \mathbf{a}_2	0,0479	0,026	0,0087	0,0446	0,1073	0,0305
\mathbf{a}_2 \mathbf{a}_3	-0,0006	0,011	-0,0205	-0,0013	0,0223	-0,0032
\mathbf{a}_3 \mathbf{a}_3	0,0154	0,009	0,0032	0,0136	0,0374	0,0052
\mathbf{g}_1 \mathbf{g}_1	0,4807	0,119	0,2431	0,4853	0,6987	0,4732
\mathbf{g}_1 \mathbf{g}_2	-0,0123	0,037	-0,0827	-0,0128	0,0603	-0,0076
\mathbf{g}_1 \mathbf{g}_3	-0,0517	0,023	-0,0977	-0,0521	-0,0039	-0,0470
\mathbf{g}_1 \mathbf{g}_4	0,0678	0,017	0,0359	0,0674	0,1023	0,0672
\mathbf{g}_1 \mathbf{g}_5	0,0004	0,011	-0,0227	0,0004	0,0232	-0,0010
\mathbf{g}_2 \mathbf{g}_2	0,0904	0,020	0,0492	0,0901	0,1334	0,0906
\mathbf{g}_2 \mathbf{g}_3	0,0262	0,010	0,0059	0,0260	0,0474	0,0253
\mathbf{g}_2 \mathbf{g}_4	-0,0213	0,008	-0,0373	-0,0211	-0,0062	-0,0202
\mathbf{g}_2 \mathbf{g}_5	-0,0004	0,005	-0,0108	-0,0003	0,0100	-0,0008
\mathbf{g}_3 \mathbf{g}_3	0,0574	0,010	0,0380	0,0571	0,0785	0,0583
\mathbf{g}_3 \mathbf{g}_4	-0,0064	0,006	-0,0173	-0,0064	0,0045	-0,0065
\mathbf{g}_3 \mathbf{g}_5	-0,0150	0,004	-0,0230	-0,0149	-0,0075	-0,0145
\mathbf{g}_4 \mathbf{g}_4	0,0473	0,006	0,0360	0,0469	0,0609	0,0442
\mathbf{g}_4 \mathbf{g}_5	-0,0068	0,004	-0,0137	-0,0068	-0,0001	-0,0069
\mathbf{g}_5 \mathbf{g}_5	0,0216	0,003	0,0155	0,0214	0,0289	0,0197
e_1 e_1	0,2587	0,014	0,2324	0,2582	0,2878	0,2575
e_2 e_2	0,2016	0,009	0,1850	0,2013	0,2196	0,2010
e_3 e_3	0,1261	0,004	0,1179	0,1260	0,1348	0,1260
e_4 e_4	0,1417	0,003	0,1357	0,1417	0,1480	0,1417

componentes. Assim como verificado para as variâncias genéticas, as trajetórias obtidas pelas metodologias AG-RA e REML-RA foram bastante semelhantes, porém todas as estimativas geradas pelo AG-RA foram levemente superiores às encontradas pelo REML-RA. As variâncias fenotípicas, de modo geral, aumentaram continuamente até a 10^a semana, decrescendo em seguida até a 30^a semana. Esse comportamento foi semelhante ao obtido para as variâncias de ambiente permanente, inclusive quanto ao aumento repentino a partir da 35^a semana de lactação.

As variâncias genéticas obtidas pelo AG-RA variaram de 0,17 a 0,29 kg², sendo levemente superiores às obtidas pelo REML-RA (que variou de 0,07 a 0,22 kg²) em toda sua trajetória. Os valores mais baixos foram observados no final da curva e os mais elevados, no início, indicando menor influência de origem genética na variância total ao final da lactação.

Com base em análise gráfica, constatou-se que as estimativas das variâncias de ambiente permanente e residuais

praticamente não diferiram entre as metodologias AG-RA e REML-RA. As estimativas das variâncias de ambiente permanente obtidas pelo AG-RA variaram de 0,22 a 0,64 kg². De modo geral, as variâncias de ambiente permanente aumentaram continuamente até a 14^a semana, decrescendo em seguida até a 29^a semana. A partir da 35^a, porém, houve repentino aumento destas estimativas, atingindo o valor de 0,64 kg².

As estimativas das herdabilidades obtidas pelo AG-RA variaram de 0,18 (39^a semana) a 0,37 (3^a semana), enquanto, pelo REML-RA, variaram de 0,09 (39^a semana) a 0,32 (3^a semana). Os valores gerados pelo AG-RA foram levemente superiores aos gerados pelo REML-RA, sendo que a partir da 30^a semana, a diferença entre eles aumentou (a maior diferença foi de 0,09), ocorrendo maior queda na curva obtida pelo REML-RA. Os maiores valores de herdabilidade foram obtidos para as semanas 3 e 4 (0,37 pelo AG-RA e 0,32 pelo REML-RA) e os mais baixos, ao final da curva (a partir da 34^a semana), como consequência das tendências observadas para as variâncias genéticas e de ambiente permanente.

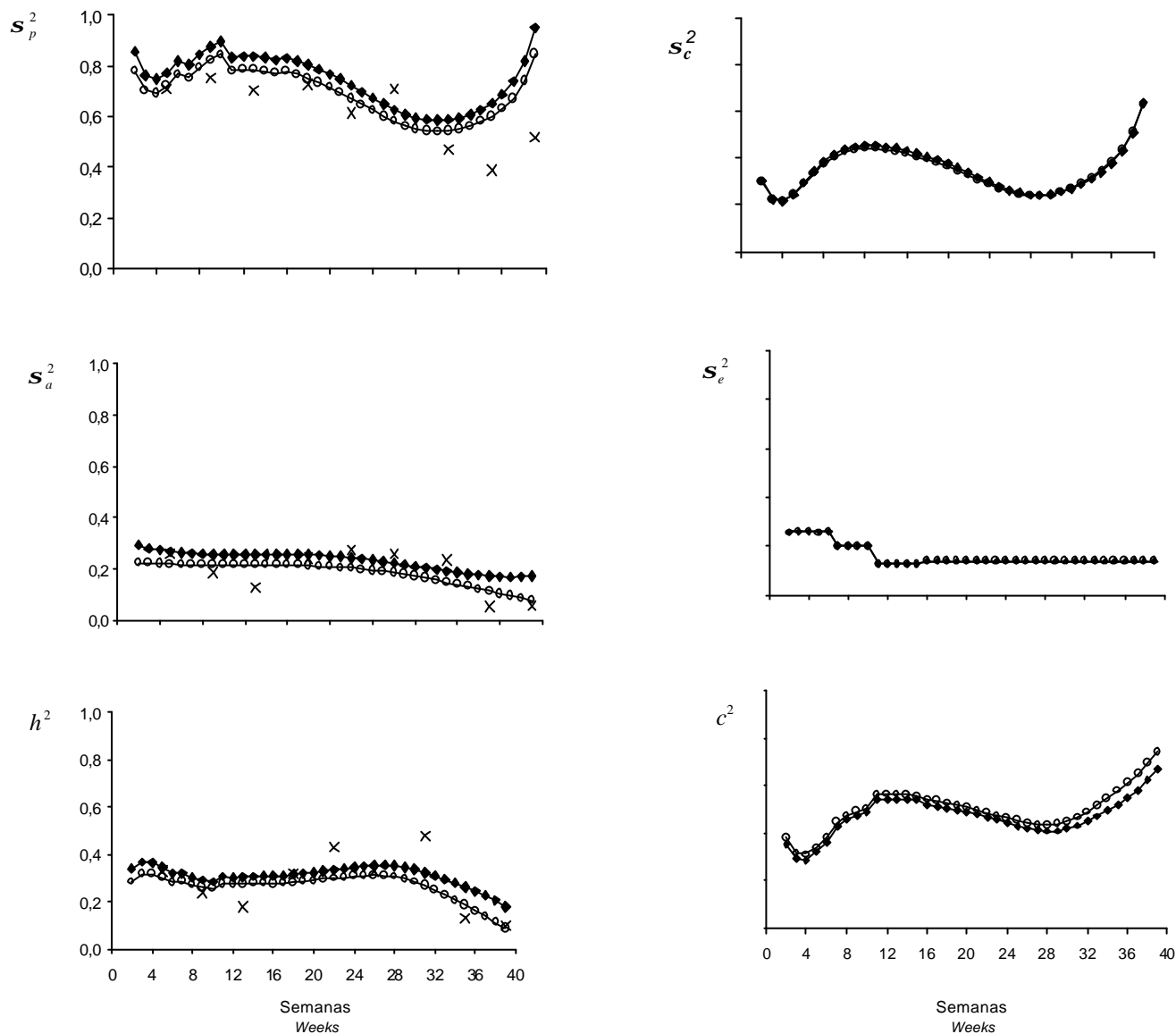


Figura 1 - Estimativas das variâncias fenotípicas (S_p^2), genéticas (S_a^2), de ambiente permanente (S_c^2) e residuais (S_e^2), das herdabilidades (h^2) e das frações da variância fenotípica referente à variância de ambiente permanente (c^2), para as PLDC obtidas para o modelo PL₃₅₄, considerando as análises AG-RA (◆), AG-UNI (×) e REML-RA (○).

Figure 1 - Estimates of phenotypic (S_p^2), genetic (S_a^2), permanent environmental (S_c^2) and residual (S_e^2) variance components, heritabilities (h^2) and fraction of phenotypic variance due to permanent environmental variance (c^2) for PLDC obtained by the PL₃₅₄ model under the AG-RA (◆), AG-UNI (×) and REML-RA (○) analyses.

No Brasil, há referências na literatura apenas da estimação de parâmetros genéticos para produção total na lactação de caprinos, não havendo relatos sobre a produção de leite no dia do controle. Além disso, estimativas de herdabilidade da produção de leite e de outras características de caprinos, a partir de número expressivo de observações, são escassas. Para um rebanho específico da raça Saanen, a herdabilidade estimada por Ribeiro (1997) foi 0,09. Gonçalves et al. (2001, 2002) estimaram a herdabilidade da produção de leite total

em caprinos das raças Alpina, Saanen e Toggenburg pelo método dos mínimos quadrados e pelo REML e encontraram valores de 0,30 e 0,22, respectivamente. Soares Filho et al. (2001), trabalhando com dois rebanhos, encontraram valores de herdabilidade iguais a 0,06; 0,01 e 0,12 para as raças Saanen, Alpina e Toggenburg, respectivamente. Pimenta Filho et al. (2004) encontraram estimativas de herdabilidade iguais a 0,23 para produção de leite total em cabras resultantes do cruzamento das raças Alpina e Gurguéia.

As estimativas de c^2 geradas por ambas as metodologias foram muito próximas, variando de 0,29 (3ª semana) a 0,67 kg² (39ª semana) pelo AG-RA e de 0,31 (4ª semana) a 0,74 kg² (39ª semana) pelo REML-RA.

As estimativas das variâncias residuais foram as mesmas para ambas as metodologias, sendo iguais a 0,26; 0,20; 0,13 e 0,14 kg² para as quatro classes estabelecidas entre as semanas 2 a 6, 7 a 10, 11 a 15 e 16 a 39, respectivamente. Nota-se, a partir das estimativas obtidas neste estudo, que a variância residual foi maior no início da lactação, decrescendo primeiramente antes do pico (atingido por volta da 8ª semana) e, logo depois, diminuindo ainda mais até o final da trajetória.

A presença de heterogeneidade de variância residual ao longo da curva de lactação tem sido relatada por diversos autores para PLDC em vacas (Rekaya et al., 2000). Essa heterogeneidade pode estar associada a diversos fatores, como estágio de prenhez, condições do parto e duração do período seco. A introdução desses efeitos nas análises muitas vezes não é viável, em decorrência da ausência de informação, sendo necessário, portanto, considerar a existência de heterogeneidade residual ao longo da curva de lactação.

As estimativas obtidas pelas análises AG-UNI para as variâncias fenotípica, genética, residual e para a herdabilidade variaram, respectivamente de 0,39 a 0,75 kg², de 0,05 a 0,27 kg², de 0,23 a 0,57 kg² e de 0,10 a 0,48. Nota-se que as trajetórias das variâncias genéticas e fenotípicas e das herdabilidades acompanham as estimativas obtidas pela análise Bayesiana univariada satisfatoriamente. As

variâncias residuais foram mais elevadas pela metodologia AG-UNI que as obtidas em ambas as análises de regressão aleatória (AG-RA e REML-RA), visto que não isolaram o efeito de ambiente permanente.

As correlações genéticas e de ambiente permanente estimadas entre as produções nos controles para o modelo PL₃₅₄ são apresentadas na Figura 2. As correlações genéticas variaram de 0,39 (estimativa obtida entre a segunda e a 39ª semanas) a 1,00 (estimativas obtidas principalmente entre duas semanas adjacentes), demonstrando que, entre controles próximos, os valores foram os maiores, decrescendo continuamente conforme o aumento da distância entre os controles. Esses resultados corroboram os encontrados nas análises de regressão aleatória para PLDC em vacas (Jensen, 2001).

As correlações genéticas obtidas neste estudo foram todas positivas, mesmo entre produções nos extremos da curva de lactação. A correlação genética entre a segunda e a 34ª semanas de lactação foi igual a 0,51, sugerindo que a seleção para maiores produções no início da lactação pode produzir efeitos positivos na produção ao final da curva (Olori et al., 1999).

As correlações de ambiente permanente variaram de 0,12 (estimativa obtida entre a segunda e a 37ª semanas) a 1,00 (estimativas obtidas entre duas semanas adjacentes, entre a 7ª e 8ª semanas até entre a 25ª e 26ª semanas).

À exceção das correlações observadas para a segunda semana do controle, cujas estimativas inicialmente decresceram rapidamente, mantendo um padrão oscilatório

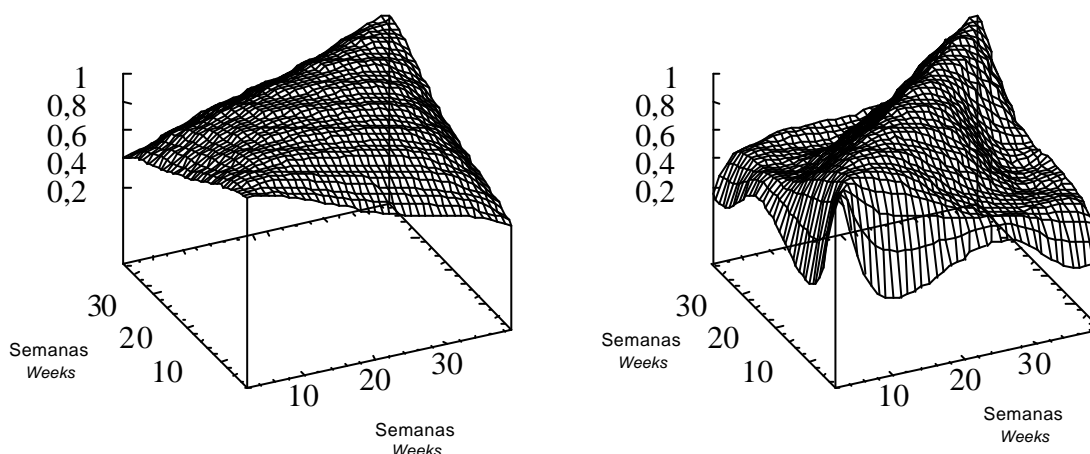


Figura 2 - Correlações genéticas (à esquerda) e de ambiente permanente (à direita) entre as produções de leite semanais, de acordo com o modelo PL₃₅₄.

Figure 2 - Genetic (left side) and permanent environmental (right side) correlations between weekly milk yields obtained by the PL₃₅₄ model.

posteriormente, as demais apresentaram decréscimo contínuo quanto maior o período entre as semanas do controle. De modo geral, as estimativas das correlações de ambiente permanente foram de menor magnitude que as genéticas.

Conclusões

A estrutura de covariâncias da PLDC em caprinos ao longo da lactação pode ser modelada adequadamente por meio da regressão aleatória.

A análise Bayesiana de regressão aleatória detectou variabilidade genética para PLDC no rebanho em estudo, o que viabiliza a predição de ganhos genéticos e o uso da seleção para indicar animais geneticamente superiores.

As estimativas das herdabilidades pela análise Bayesiana de regressão aleatória foram de magnitude moderada ao longo de toda a trajetória da lactação, sugerindo que a PLDC pode ser recomendada para uso no melhoramento genético de caprinos.

Os resultados gerados nas análises de regressão aleatória pela metodologia Bayesiana e pelo REML foram semelhantes.

Agradecimento

À CAPES, pela bolsa de estudos concedida à primeira, ao terceiro e quarto autores durante seus respectivos cursos de doutoramento.

À FAPEMIG, pela bolsa de pesquisadora visitante concedida à segunda autora.

Literatura Citada

- BROTHESTONE, S.; WHITE, I.M.S.; MEYER, K. Genetic modeling of daily milk yield using orthogonal polynomials and parametric curves. **Animal Science**, v.70, p.407-415, 2000.
- FAO. "Chapter on Brazil", 2000. In: FAO, AGRICULTURE, TRADE AND FOOD SECURITY. Disponível em: <http://www.fao.org> Acesso em: dezembro de 2004.
- GONÇALVES, H.C.; SILVA, M.A.; WECHSLER, F.S. et al. Fatores genéticos e de meio na produção de leite de caprinos leiteiros. **Revista Brasileira de Zootecnia**, v.30, n.3, p.719-729, 2001.
- 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.
- GEMAN, S., GEMAN, D. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v.6, p.721-741, 1984.
- GIANOLA, D.; FERNANDO, R.L. Bayesian methods in animal breeding theory. **Journal of Animal Science**, v.63, p.217-244, 1986.
- MADSEN, P.; JENSEN, J.; THOMPSON, R. 1994. Estimation of (co)variance components by REML in multivariate mixed linear models using average of observed and expected information. In: WORLD CONGRESS OF GENETICS APPLIED TO LIVESTOCK PRODUCTION, 5., 1994, Guelph. **Proceedings...** Guelph: 1994. p.19-22.
- JAMROZIK, J.; SCHAEFFER, L.R. Estimates of genetic parameters for a test day model with random regressions for yield traits of first lactation Holsteins. **Journal of Dairy Science**, v.80, n.4, p.762-770, 1997.
- JAMROZIK, J.; SCHAEFFER, L.R. Bayesian comparison of random regression models for test-day yields in dairy cattle. In: WORLD CONGRESS OF GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier. **Proceedings...** Montpellier: 2002 (CD-ROM).
- JENSEN, J. Genetic evaluation of dairy cattle using test-day models. **Journal of Dairy Science**, v.84, p.2803-2812, 2001.
- KIRKPATRICK, M.; LOFSVOLD, D.; BULMER, M. Analysis of the inheritance, selection and evolution of growth trajectories. **Genetics**, v.124, p.979-993, 1990.
- 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).
- MEYER, K. "RRGIBBS" – A program for simple random regression analyses via Gibbs Sampling. In: WORLD CONGRESS OF GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7., 2002, Montpellier. **Proceedings...** Montpellier: 2002. CD-ROM.
- MISZTAL, I.; STRABEL, T.; JAMROZIK, J. et al. Strategies for estimating the parameters needed for different test-day models, **Journal of Dairy Science**, v.83, p.1125-1134, 2000.
- OLORI, V.E.; HILL, W.G.; MCGUIRK, S. et al. Estimating variance components for test day milk records by restricted maximum likelihood with random regression animal model. **Livestock Production Science**, v.61, p.53-63, 1999.
- 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.
- REKAYA, R.; CARABAÑO, M.J.; TORO, M.A. Use of test day yields for the genetic evaluation of production traits in Holstein-Friesian cattle. **Livestock Production Science**, v.57, p.203-217, 1999.
- REKAYA, 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.
- RIBEIRO, A.C. **Estudo dos efeitos genéticos e de ambiente sobre características de importância econômica em caprinos da raça Saanen**. Jaboticabal: Universidade Estadual Paulista, 1997. 119p. Dissertação (Mestrado em Zootecnia) - Universidade Estadual Paulista, 1997.
- SOARES FILHO, G.; MCMANUS, C.; MARIANTE, A.S. Fatores genéticos e ambientais que influenciam algumas características de reprodução e produção de leite em cabras no Distrito Federal. **Revista Brasileira de Zootecnia**, v.30, n.1, p.133-140, 2001.
- SORENSEN, D.; GIANOLA, D. **Likelihood, Bayesian, and MCMC methods in quantitative genetics**. New York: Springer, 2002. 740p.
- SOUZA, G.N.; MOREIRA, E.C.; RISTOW, P. et al. Formas de exploração do rebanho caprino no Estado do Rio de Janeiro, 1998/2000. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.54, n.2, 2002 (Comunicação).
- Van KAAM, J.B.C.H.M. "GIBANAL" – Analyzing program for Markov Chain Monte Carlo Sequences (Version 2.10). Manual, Department of Animal Sciences, Wageningen Agricultural University, 1998. não-paginado.
- Van TASSELL, C.P.; Van VLECK, L.D. **A manual for use of MTGSAM. A set of Fortran programs to apply Gibbs Sampling to animal models for variance components estimation**. Lincoln: Department of Agriculture, Agricultural Research Service, 1995. 85p.