

ABORDAGEM BAYESIANA E FREQUENTISTA EM ANÁLISE GENÉTICA TRICARÁTER PARA CRESCIMENTO E REPRODUÇÃO DE BOVINOS NELORE

CARINA UBIRAJARA DE FARIA,¹ CLÁUDIO DE ULHÔA MAGNABOSCO,² LUCIA GALVÃO ALBUQUERQUE,³ ARCADIO DE LOS REYES,⁴ LUIZ ANTÔNIO FRAMARTINO BEZERRA⁵ E RAYSILDO BARBOSA LOBO⁶

1. Professora adjunto da Universidade Federal de Goiás, Campus de Jataí, Jataí, GO.

2. Pesquisador da Embrapa Cerrados, bolsista CNPq, Planaltina, DF, Brasil.

3. Professora adjunto do Departamento de Melhoramento Animal, FCAV/UNESP, Jaboticabal, SP, Brasil.

4. Professor titular do Departamento de Produção Animal, EV/UFG, Goiânia, GO, Brasil

5. Pesquisador do Departamento de Genética da FMRP-USP, Ribeirão Preto, SP, Brasil.

6. Professor associado do Departamento de Genética da FMRP-USP, Ribeirão Preto, SP, Brasil.

RESUMO

Os objetivos deste estudo foram estimar os componentes de variância e parâmetros genéticos de características de crescimento e reprodução de bovinos da raça Nelore, em modelo animal tricaráter, utilizando a Máxima Verossimilhança Restrita e a Inferência Bayesiana. Na implementação do Amostragem de Gibbs, que aplica a teoria bayesiana, consideraram-se os três diferentes níveis de conhecimento das informações iniciais dos componentes de variância e covariância. O conjunto de dados continha 15.173 registros de peso aos 365 dias de idade (P365), 6.911 registros de perímetro escrotal aos 365 dias de idade (PE365) e 10.388 dados de idade ao primeiro parto (IPP). O

modelo linear incluiu o grupo de contemporâneos e a idade da vaca ao parto como efeitos fixos (exceto para IPP), e os efeitos genético aditivo direto e residual como aleatórios. As estimativas obtidas pela Máxima Verossimilhança Restrita foram diferentes das obtidas pela Amostragem de Gibbs. No entanto, considerando os três diferentes níveis de informação inicial, não houve diferenças nas estimativas dos parâmetros genéticos obtidas pela Amostragem de Gibbs. O método bayesiano apresentou vantagens em relação ao método frequentista, graças às distribuições marginais posteriores, que ofereceram mais informações sobre os parâmetros estimados.

PALAVRAS-CHAVES: Bovinos de corte, modelo tricaráter, parâmetros genéticos.

ABSTRACT

BAYESIAN AND FREQUENTIST APPROACHES IN THREE TRAIT GENETIC ANALYSIS FOR GROWTH AND REPRODUCTION IN NELORE CATTLE

The Bayesian inference has been proposed as an alternative for estimating variance components instead of the frequentist approach. The objective of this study was to estimate the (co)variance components and genetic parameters of growth and reproduction traits in Nelore cattle using three-trait animal model by Restricted Maximum Likelihood and Gibbs Sampling. The data set had 15,173 and 6,911 records of weight (W365) and scrotal

circumference (SC365) at 365 days of age, respectively, and 10,388 records of age at first calving (AFC), from the Nelore Breed Genetic Improvement Program farms. The linear model included the contemporary groups and age-of-dam (with exception for AFC) as fixed effects, and residual and additive direct genetic effects as random. Estimates obtained by Restricted Maximum Likelihood were different from those found by Gibbs Sampling. There

were no differences in genetic parameter estimates using the three levels of prior information by Gibbs Sampling. The Bayesian inference has advantages in relation to

Frequentist approach due to marginal distributions which offer more information about the parameters.

KEY WORDS: Beef cattle, genetic parameters, three-trait models.

INTRODUÇÃO

O aprimoramento de métodos de estimação de componentes de variância tem sido cada vez mais importante na pecuária de corte. Com a abertura de novos mercados e o crescimento acelerado da indústria da carne, a utilização de valores genéticos na seleção de reprodutores e matrizes tornou-se uma prática comum no Brasil. No entanto, os componentes de variância, necessários para a obtenção dos valores genéticos, precisam ser estimados de forma acurada para que as diferenças entre o valor predito e o verdadeiro sejam minimizadas.

Nesse contexto, vários métodos de estimação de componentes de variância têm sido utilizados. Atualmente, a Máxima Verossimilhança Restrita (REML) é o método de escolha pelos programas de melhoramento genético no Brasil e seu uso foi intensificado pela disponibilidade do programa Multiple Trait Derivate-Free Restricted Maximum Likelihood – MTDFREML (BOLDMAN et al., 1995). No entanto, procedimentos bayesianos têm sido propostos como uma alternativa de aplicação em relação aos métodos frequentistas nas análises genéticas considerando o modelo de touro (WANG et al., 1993), o modelo animal (VAN TASSEL & VAN VLECK, 1996), o modelo de limiar para dados categóricos (VAN TASSEL et al., 1998) e na obtenção da resposta à seleção (SORENSEN et al., 1994). No Brasil, o primeiro estudo que utilizou a inferência bayesiana na estimação de componentes de variância, para características produtivas em bovinos, foi realizado por MAGNABOSCO (1997). A aplicação dos métodos de Markov Chain Monte Carlo (MCMC), dentre os quais se destaca a Amostragem de Gibbs (GS), é o que propicia a inferência bayesiana. A Amostragem de Gibbs é aplicada para gerar um valor para cada parâmetro desconhecido e apresenta fácil implementação, principalmente quando comparada a algoritmos

baseados em processos não derivativos, uma vez que os resultados permitem uma inferência bayesiana gerando distribuições posteriores marginais completas, a partir das quais são obtidas as estimativas dos componentes de (co)variâncias e parâmetros genéticos.

Outro fator importante nas análises genéticas é a definição correta do modelo linear. Nesse contexto, é comum se dispor de mais de uma característica mensurada para um único indivíduo, sendo necessária a utilização de modelos lineares multicaráteres para estimação dos componentes de variância e predição dos valores genéticos dos indivíduos. As análises multicaráteres permitem explorar a correlação entre as características, aumentando a acurácia da predição genética, além de serem vantajosas, por incluírem informações de dados referentes à característica em que a seleção foi pequena ou nula. De acordo com MARTINS et al. (1997), um aspecto indesejável em análises multicaráteres é que o número de elementos a serem preditos torna-se excessivamente grande quando se envolvem todos os indivíduos identificados na estrutura genealógica da população. Nas análises genéticas tricaráteres, o número de variáveis aleatórias genéticas e residuais a serem estimadas é grande, e o número de equações a serem resolvidas equivale ao número de animais, multiplicado pelo número de características; assim, a aplicação de um método robusto e consistente é indispensável.

Este trabalho teve o objetivo de estimar os componentes de (co)variâncias e parâmetros genéticos de características de crescimento e reprodução de bovinos da raça Nelore, em modelo animal tricaráter, utilizando os métodos da Máxima Verossimilhança Restrita e Amostragem de Gibbs.

MATERIAL E MÉTODOS

Os dados utilizados foram provenientes de fazendas participantes do Programa de Melho-

ramento Genético da Raça Nelore (PMGRN) e cedidos pela Associação Nacional de Criadores e Pesquisadores (ANCP). O conjunto de dados continha 15.173 registros de peso aos 365 dias de idade (P365), 6.911 registros de perímetro escrotal aos 365 dias de idade (PE365) e 10.388 dados de idade ao primeiro parto (IPP). Todos os machos que apresentaram medidas para PE365 tam-

bém possuíam registros para P365, no entanto, nenhuma fêmea com informação para IPP possuía registro para P365. Assim, para o início das análises genéticas, se dispôs de dados de campo referentes a 25.561 animais e matriz de parentesco de 50.114 animais. A estatística descritiva dos dados é apresentada na Tabela 1.

TABELA 1. Estatística descritiva dos dados para as características peso e perímetro escrotal aos 365 dias, e idade ao primeiro parto de bovinos da raça Nelore.

Característica	Média	DP	CV (%)	Mínimo	Máximo
Peso (P365)	233,84	35,16	15,04	93,00	378,00
Perímetro escrotal (PE365)	19,38	2,05	10,57	11,00	29,00
Idade ao primeiro parto (IPP)	35,77	3,48	9,73	24,00	44,00

DP_ desvio-padrão; CV= coeficiente de variação.

Os grupos contemporâneos (GC) consistiram de fazenda, ano, estação de nascimento e lotes de manejo para o perímetro escrotal e idade ao primeiro parto. Para peso aos 365 dias de idade foi incluído o sexo na formação dos grupos contemporâneos. Considerou-se o efeito da idade da vaca ao parto (IVP) como efeito fixo nas análises, com exceção da idade ao primeiro parto. Para a estimação dos componentes de (co) variâncias e parâmetros genéticos utilizaram-se dois métodos de análise: Máxima Verossimilhança Restrita e Amostragem de Gibbs. O modelo matricial incluiu efeitos fixos de grupos contemporâneos e de idade da vaca ao parto, e as contribuições dos efeitos genéticos aditivos diretos e residuais como aleatórios, conforme descrito a seguir:

$$y = X\beta + Za + e$$

em que y é o vetor das variáveis dependentes (P365, PE365 e IPP), β o vetor dos efeitos fixos (GC e IVP), X a matriz de incidência que associa β com y , a representa o vetor dos efeitos genéticos aditivos diretos, Z a matriz de incidência que associa a com y e, finalmente, e o vetor de efeitos residuais. Assumiu-se que y , a , e possuem distribuição conjunta normal multivariada, como segue:

$$\begin{bmatrix} y \\ a \\ e \end{bmatrix} \sim NMV \left\{ \begin{bmatrix} X\beta \\ \phi \\ \phi \end{bmatrix}, \begin{bmatrix} ZGZ' + R & ZG & R \\ GZ' & G & \phi \\ R & \phi & R \end{bmatrix} \right\}$$

em que $G = As_a^2$ e $R = Is_e^2$, sendo A a matriz de coeficientes de parentesco, I a matriz de identidade, s_a^2 a variância genética aditiva e s_e^2 a variância residual. No modelo tricaráter, $G = G_0 \otimes A$ e $R = I \otimes R_0$, e as pressuposições usuais são apresentadas como:

$$G_0 = \begin{bmatrix} \sigma_{a_1}^2 & \sigma_{a_1 a_2} & \sigma_{a_1 a_3} \\ \sigma_{a_1 a_2} & \sigma_{a_2}^2 & \sigma_{a_2 a_3} \\ \sigma_{a_1 a_3} & \sigma_{a_2 a_3} & \sigma_{a_3}^2 \end{bmatrix}, R_0 = \begin{bmatrix} \sigma_{e_1}^2 & \sigma_{e_1 e_2} & 0 \\ \sigma_{e_1 e_2} & \sigma_{e_2}^2 & 0 \\ 0 & 0 & \sigma_{e_3}^2 \end{bmatrix}$$

em que G_0 é a matriz de variâncias e covariâncias genética aditiva para as três características, \otimes é o operador do produto direto, e R_0 é a matriz de variâncias e covariâncias residuais para as características mensuradas no animal.

Considerando o método frequentista, obtiveram-se as estimativas dos componentes de (co)variâncias e parâmetros genéticos por meio do programa MTDFREML, descrito por BOLDMAN et al. (1995). Adotou-se, como critério de convergência, que a variância de $-2\log$ (valor

da função de verossimilhança) do simplex fosse menor que 10^{-9} . Foram necessários seis reinícios, utilizando-se os resultados da rodada anterior como valores iniciais para garantir a convergência a um máximo global.

Para a obtenção das estimativas dos componentes de (co)variâncias e parâmetros genéticos pelo procedimento bayesiano, foi utilizado o programa Multiple Trait using Gibbs Sampling under Animal Model (MTGSAM) desenvolvido por VAN TASSEL & VAN VLECK (1996). Para as distribuições iniciais das variâncias e covariâncias, assumiu-se uma distribuição Wishart Invertida (WI) para os efeitos genéticos e residuais de todas as características, inclusive as covariâncias entre elas. Se a variável aleatória, W , é uma variável WI, então a função densidade de probabilidade de W é:

$$p(W|v, V) \propto K(v, V^{-1}) |W|^{-\frac{1}{2}(v+m+1)} \exp\left(\text{tr}\left(-\frac{1}{2}V^{-1}W^{-1}\right)\right)$$

$W, V > 0$; $v > m + 1$, em que:

$$K(v, V^{-1}) = |V^{-1}|^{\frac{1}{2}v} \left(2^{\frac{1}{2}vm} \pi^{\frac{1}{4}m(m-1)} \prod_{i=1}^m \Gamma\left(\frac{v-j+1}{2}\right) \right)^{-1}$$

Conforme descrito por JOHNSON & KOTZ (1972), são necessários diversos parâmetros para a função densidade de probabilidade das variáveis aleatórias. O parâmetro v é o grau de liberdade correspondente à variável WI e ao grau de confiabilidade da distribuição inicial. A matriz V descreve a estrutura de variâncias e covariâncias da variável W , e m é a dimensão de V . A média de W é V^{-1}/v^* , em que $V^{-1} = v^*V_0$, V_0 é a matriz de variâncias e covariâncias especificada pela informação inicial.

As distribuições condicionais completas dos parâmetros genéticos foram obtidas das densidades posteriores conjuntas, ou seja, das densidades dos parâmetros, considerando as observações e as informações iniciais. No contexto da inferência bayesiana, as informações *a priori* sobre os parâmetros são utilizadas em associação com os dados amostrais pela função de verossimilhança, gerando, assim, uma distribuição conjunta *a posteriori* (Van TASSEL & VAN VLECK, 1996;

BLASCO, 2001; GUEDES et al., 2005; MELLO et al., 2006). Dessa forma, as distribuições marginais dos parâmetros, obtidas a partir da integração da distribuição conjunta *a posteriori*, fornecem os estimadores (média, moda e mediana) de interesse.

O *software* MTGSAM utiliza a iteração de Gauss Seidel para obter os valores iniciais dos efeitos fixos e aleatórios, baseados nos valores iniciais dos componentes de (co)variância. Neste estudo, a matriz inicial de variâncias e covariâncias considerou três níveis de conhecimento das informações iniciais: não-informativa ($v = 0$), pouco informativa ($v = 5$) e informativa ($v = 9$). No contexto bayesiano, é possível distinguir diferentes níveis de conhecimento sobre a informação inicial dos componentes de (co)variância: (1) valores iniciais não-informativos implicam que os valores são definidos arbitrariamente, sem nenhum conhecimento prévio; (2) valores iniciais pouco informativos indicam que existe algum conhecimento prévio; e (3) os valores iniciais informativos indicam que a informação inicial é estabelecida mais claramente, com conhecimento prévio. Cada um desses três níveis de conhecimento da informação inicial foi modelado ajustando o grau de conhecimento do parâmetro v .

Na implementação do Algoritmo de Gibbs, não é possível considerar que as covariâncias sejam restritas a zero, como acontece no método REML, em que determinadas covariâncias podem ser fixadas em zero. Conseqüentemente, os efeitos aleatórios não correlacionados e residuais precisam ser agrupados por características para impor zero nas covariâncias entre efeitos que não poderiam ser estimados (Van TASSEL & Van VLECK, 1996). Essa subdivisão dos efeitos aleatórios e residuais não-correlacionados é empregada, principalmente, para características limitadas ao sexo. Deve-se ressaltar que, para designar covariância residual como valor zero, entre duas características, é preciso que essas sejam consideradas como membros de diferentes grupos, mas isso não implica que sejam diferentes grupos genéticos. Neste estudo, foi considerado um grupo para peso e perímetro escrotal aos 365 dias de idade, e a idade ao primeiro parto como

outro grupo, sendo estimada somente a covariância residual entre P365 e PE365.

Na implementação da Amostragem de Gibbs foram utilizadas 300.000 iterações com descarte inicial de 20.000 iterações para o período de aquecimento da cadeia de Gibbs, no intuito de se minimizarem os efeitos dos valores iniciais. Para assegurar a independência das amostras, considerou-se um intervalo de retirada de 100 iterações, gerando um total de 2.800 amostras dos componentes de variância e covariância. Procedeu-se às análises das amostras, da correlação serial e da convergência da cadeia de Gibbs com o auxílio do programa GIBANAL (VAN KAAM, 1998). O erro de Monte Carlo é o erro da estimação do parâmetro em virtude do número de amostras usadas na cadeia de Gibbs. Estimou-se o erro de Monte Carlo mediante o cálculo da variância das amostras retiradas para cada componente de (co)variância e dividindo-se essa variância pelo número de amostras. Assim, a raiz quadrada desse valor é uma aproximação do desvio-padrão do erro associado com o tamanho da cadeia de Gibbs.

RESULTADOS E DISCUSSÃO

As estimativas dos componentes de (co)variâncias e parâmetros genéticos para o peso e perímetro escrotal aos 365 dias e idade ao primeiro parto, obtidas pelos métodos REML e GS sob análise tricaráter, são apresentadas na Tabela 2. Observa-se que as estimativas obtidas pelo REML foram diferentes quando comparadas ao método GS. Considerando, em especial, os parâmetros genéticos, as estimativas de herdabilidade para P365, PE365 e IPP obtidas pelo REML foram menores do que aquelas obtidas pelo método GS, o que não ocorreu em relação às correlações genéticas, sendo que pelo método REML essas foram maiores do que pelo GS. No entanto, em ambos os métodos, os valores obtidos para os parâmetros genéticos são reportados na literatura (KOOTIS et al., 1994; MERCADANTE et al., 2000; PEREIRA et al., 2001; DIAS et al., 2000; DIAS, 2001; FARIA, 2003). Considerando o método REML, MERCADANTE et al. (2000), em

análise bicaráter, obtiveram herdabilidades de 0,35 para P365, 0,27 para IPP e correlação genética de -0,23 para rebanhos Nelore. Já no estudo realizado por DIAS et al. (2000), também em análise bicaráter e utilizando o REML, as estimativas de herdabilidade para PE365 e IPP foram de 0,43 e 0,17, respectivamente, e a correlação entre PE365 e IPP de -0,19.

Pelo método da Amostragem de Gibbs, em análise bicaráter, FARIA (2003) encontrou estimativas semelhantes às do presente estudo. As estimativas de herdabilidade de PE365 e IPP foram 0,67 e 0,41, respectivamente, e a correlação genética de -0,19; para análise bicaráter de P365 e PE365, os valores foram de 0,49 e 0,59 para herdabilidade e 0,51 para a correlação genética, respectivamente. PEREIRA et al. (2001), em análise tricaráter para peso ao sobreano, perímetro escrotal e idade ao primeiro parto na raça Nelore, encontraram herdabilidades de 0,40, 0,33 e 0,10, respectivamente, utilizando o método REML. Já as correlações genéticas se mostraram de maior magnitude que as encontradas neste estudo pelo mesmo método, com exceção da correlação genética entre peso e perímetro escrotal, que foi de 0,36.

Na Tabela 2 observa-se que as estimativas dos parâmetros genéticos para os três níveis de informação estudados foram idênticas. De acordo com BLASCO et al. (1998), as diferentes distribuições iniciais assumidas na implementação do GS podem afetar as conclusões de experimentos. Muitos estudos de aplicação do GS têm usado valores iniciais não-informativos (WANG et al., 1993; VAN TASSEL & VAN VLECK, 1996; MAGNABOSCO et al., 2000). Nesse caso, a distribuição inicial da cadeia de Gibbs é assumida como uma distribuição uniforme com a mesma probabilidade para todos os possíveis valores do parâmetro (LIN & BERGER, 2001). Ambos os procedimentos REML e GS são totalmente dependentes das informações dos dados quando utilizam valores iniciais não-informativos. No entanto, nos procedimentos bayesianos, BLASCO et al. (1998) constataram que, mesmo empregando diferentes informações iniciais sobre os parâmetros, as inferências posteriores não

diferiram substancialmente. No presente estudo, uma possível explicação para o fato de as estimativas dos parâmetros genéticos não diferirem nos três níveis de informação, *a priori*, está no número de dados para cada característica, sendo suficientes para gerarem estimativas precisas. De

acordo com HOBERT & CASELLA (1996), os valores iniciais não-informativos podem ser utilizados em situações em que há um maior número de observações das variáveis dependentes sem comprometer os resultados das análises.

TABELA 2. Estimativas dos componentes de (co)variâncias e parâmetros genéticos para peso (P365) e perímetro escrotal (PE365) aos 365 dias e idade ao primeiro parto (IPP), em análise tricaráter, de acordo com o método de estimação.

Parâmetros	Método de estimação			
	REML	GS ^{NI}	GS ^{PI}	GS ^I
(Co)variância				
P365 (σ_a^2)	265,51	341,74±0,490	340,35±0,490	342,25±0,500
P365PE365 (σ_a)	10,55	13,19±0,030	13,16±0,031	13,23±0,031
P365IPP (σ_a)	-3,37	-2,65±0,037	-2,69±0,037	-2,72±0,037
PE365 (σ_a^2)	1,46	1,86±0,003	1,86±0,003	1,88±0,003
PE365IPP (σ_a)	-0,55	-0,48±0,003	-0,49±0,003	-0,50±0,003
IPP (σ_a^2)	2,13	3,42±0,006	3,39±0,006	3,39±0,006
P365 (σ_c^2)	412,69	362,64±0,350	363,24±0,350	361,72±0,350
P365PE365 (σ_c)	14,58	12,86±0,022	12,87±0,022	12,81±0,022
PE365 (σ_c^2)	1,50	1,25±0,002	1,24±0,002	1,22±0,002
IPP (σ_c^2)	5,85	4,97±0,004	4,99±0,004	4,98±0,004
Parâmetro genético				
P365 (h^2)	0,39	0,48±0,0006	0,48±0,0006	0,48±0,0006
PE365 (h^2)	0,49	0,60±0,0007	0,60±0,0007	0,60±0,0007
IPP (h^2)	0,27	0,40±0,0006	0,40±0,0006	0,40±0,0006
P365PE365 (r_g)	0,54	0,52±0,0008	0,52±0,0008	0,52±0,0008
P365IPP (r_g)	-0,14	-0,08±0,0011	-0,08±0,0011	-0,08±0,0011
PE365IPP (r_g)	-0,31	-0,19±0,0011	-0,19±0,0012	-0,19±0,0011

σ_a^2 : variância genética aditiva, σ_a : covariância genética aditiva, σ_c^2 : variância residual, h^2 : herdabilidade, r_g : correlação genética, GS: médias posteriores e erros padrão, NI: não informativa, PI: pouco informativa e I: informativa.

As médias e os intervalos de credibilidade dos parâmetros genéticos obtidos a partir das 2.800 amostras da cadeia de Gibbs são apresentados na Figura 1. De acordo com o gráfico de dispersão, podem-se comparar as estimativas pontuais do REML com a média da distribuição marginal posterior obtida pelo GS. Observa-se que as estimativas dos parâmetros genéticos obtidas pelo método REML foram menores, localizando-se fora da região de credibilidade das estimativas

obtidas pelo método GS, exceto para a correlação genética entre peso e perímetro escrotal aos 365 dias de idade. No entanto, a variação entre a estimativa mínima e máxima da região de credibilidade foi pequena, indicando que a convergência da cadeia de Gibbs foi alcançada. De acordo com SILVA et al. (2003), a amplitude entre o valor máximo e mínimo do intervalo de confiança permite inferir, com maior confiabilidade, as estimativas obtidas. Como as estimativas obtidas pelo

GS, considerando os três níveis de informação *a priori*, foram muito similares para os parâmetros genéticos, não houve a necessidade de representar graficamente a dispersão das estimativas para as três análises.

LIN & BERGER (2001) avaliaram o comportamento das distribuições marginais posteriores dos três níveis de informação *a priori* utilizando dados de pequenas populações. Segundo os autores, as estimativas foram distintas quando os valores iniciais eram informativos, e a distribuição posterior dos parâmetros era mais simétrica e com menor variação. Destacaram, porém, que a informação proveniente dos dados foi insuficiente, o que favoreceu a maior variação quando os valores iniciais eram não-informativos, uma vez que, neste caso, as estimativas dependem das informações dos dados. Segundo VAN TASSEL & VAN VLECK (1996), a obtenção de distribuições marginais posteriores inadequadas com o

uso de informação inicial desconhecida não tem sido bem caracterizada em análises multicarateres, favorecendo esse tipo de análise.

GIANOLA & FOULLEY (1990) sugeriram que todos os componentes de (co)variâncias presentes em um modelo não são estimados com a mesma acurácia. Na Tabela 3, são apresentados os coeficientes de variação para as estimativas dos componentes de (co)variâncias obtidos pelo GS nas três diferentes análises. Observa-se que, para o peso aos 365 dias e idade ao primeiro parto, a variância residual foi estimada com menor variação relativa à média, e a variância genética mostrou maior variação. As covariâncias genéticas foram estimadas com menor precisão do que as variâncias genéticas. Em relação ao uso de diferentes níveis de conhecimento dos valores iniciais, os coeficientes de variação não apresentaram grandes diferenças.

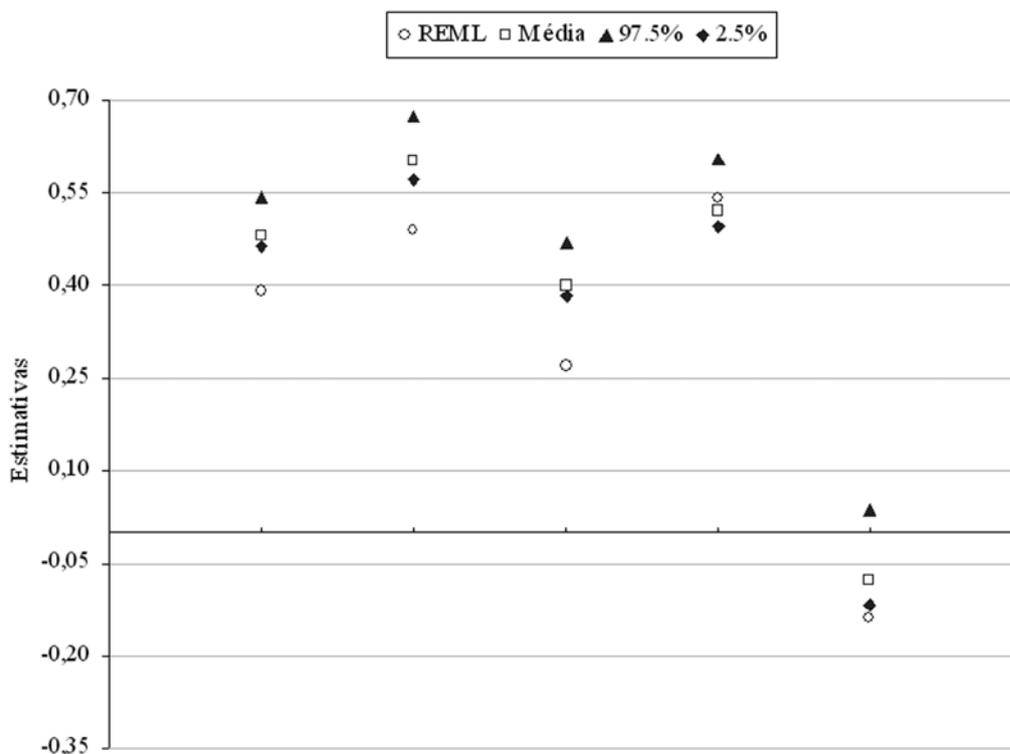


FIGURA 1. Estimativas de parâmetros genéticos para peso (P_{365}) e perímetro escrotal (PE_{365}) aos 365 dias e idade ao primeiro parto (IPP_3), em análise tricaráter, considerando os métodos da Máxima Verossimilhança Restrita e Amostragem de Gibbs (região de credibilidade a 95%).

Na Tabela 3 também é apresentado o desvio-padrão do erro de Monte Carlo da análise tricaráter, considerando os três níveis de informação *a priori* dos valores iniciais. De acordo com VAN TASSEL & VAN VLECK (1996), o erro de Monte Carlo é inversamente proporcional ao tamanho da cadeia de Gibbs e conhecer esse fator é muito importante para avaliar se a implementação da Amostragem de Gibbs foi

adequada para gerar as médias posteriores das distribuições marginais dos componentes de (co) variâncias. Neste estudo, observa-se que o erro de Monte Carlo foi praticamente nulo para todos os componentes de (co)variâncias, indicando que o tamanho da cadeia de Gibbs foi suficiente para obter estimativas precisas das médias posteriores e que a convergência foi alcançada.

TABELA 3. Coeficiente de variação das amostras e desvio-padrão (DP) do erro de Monte Carlo das estimativas dos componentes de (co)variâncias obtidos pelo método da Amostragem de Gibbs em análise tricaráter utilizando três níveis de informação *a priori*.

(Co)variâncias	Coeficiente de variação			Erro de Monte Carlo ^{DP}		
	NI	PI	I	NI	PI	I
P365 (σ_a^2)	7,63	7,70	7,67	0,493	0,492	0,496
P365 (σ_e^2)	5,07	5,09	5,11	0,347	0,349	0,350
PE365 (σ_a^2)	7,98	8,08	8,07	0,003	0,003	0,003
PE365 (σ_e^2)	8,34	8,52	8,72	0,002	0,002	0,002
IPP (σ_a^2)	9,13	9,26	9,24	0,006	0,006	0,006
IPP (σ_e^2)	4,57	4,58	4,58	0,004	0,004	0,004
P365PE365 (σ_a)	12,27	12,43	12,43	0,031	0,031	0,031
P365PE365 (σ_e)	8,94	9,02	9,11	0,022	0,022	0,022
P365IPP (σ_a)	-74,40	-73,59	-72,57	0,037	0,037	0,037
PE365IPP (σ_a)	-31,73	-31,46	-30,90	0,003	0,003	0,003

NI: não-informativa PI: pouco informativa I: informativa. σ_a^2 : variância genética aditiva, σ_e^2 : variância residual, σ_a : covariância genética, σ_e : covariância residual

Na Figura 2 são apresentadas as densidades marginais posteriores das variâncias genéticas obtidas a partir de três níveis de informação *a priori* da distribuição inicial dos componentes de (co)variâncias. Mostram-se apenas as densidades marginais posteriores das covariâncias genéticas, pelo fato de estas terem apresentado maiores variações entre as amostras. A moda das distribuições marginais posteriores obtidas pelo GS não corresponde às estimativas obtidas pelo REML. De acordo com GIANOLA & FOULLEY (1990), uma possível limitação do método REML seria que as estimativas são elementos do vetor modal da distribuição posterior conjunta dos componentes de

variância, sendo que a moda da distribuição marginal de cada componente poderia fornecer uma melhor aproximação da média, do que um componente da moda conjunta. Pode-se observar que as distribuições posteriores foram estáveis e com tendência a uma distribuição normal, de modo que tanto as medidas de tendência central como a média, moda e mediana apresentaram similaridade. RESENDE (2000) relata que os procedimentos bayesianos propiciam estimativas mais precisas e que a grande vantagem refere-se à obtenção dos desvios-padrão e intervalo de confiança para os parâmetros genéticos, bem como à obtenção das densidades marginais posteriores.

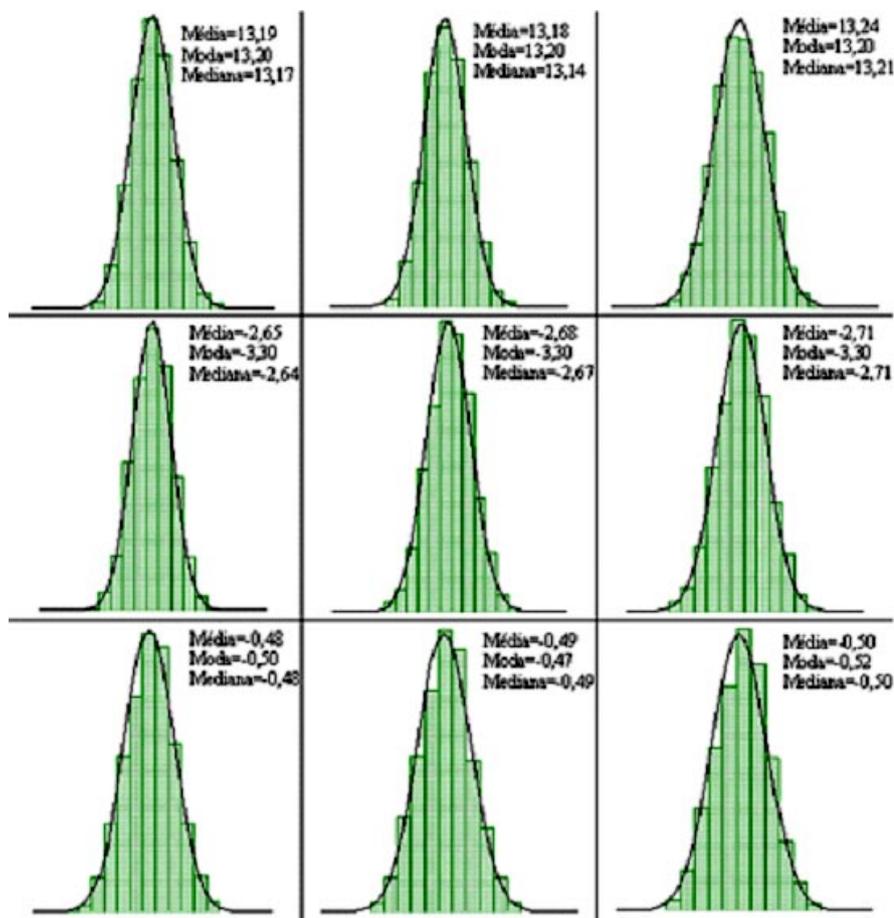


FIGURA 2. Densidades marginais posteriores das estimativas de covariâncias genéticas para P365 e PE365 (superior), P365 e IPP (central) e PE365 e IPP (inferior), em análise tricaráter pelo método da Amostragem de Gibbs, utilizando três níveis de informação *a priori*: não-informativa (esquerda), pouco informativa (centro) e informativa (direita).

CONCLUSÕES

O método bayesiano, mediante a utilização da Amostragem de Gibbs, permitiu a obtenção de estimativas pontuais e intervalares dos parâmetros genéticos. Assim, a abordagem bayesiana apresentou vantagens em relação ao método frequentista, graças às distribuições marginais posteriores, que ofereceram mais informações sobre os parâmetros genéticos do modelo tricaráter.

REFERÊNCIAS

- BOLDMAN, K.G.; KRIESE, L.A.; VAN VLECK, L.D.; VAN TASSEL, C.P.; KACHMAN, S.D. **A manual for use of MTDFREML. A set of programs to obtain estimates of variances and covariances (DRAFT).** Lincoln: Department of Agriculture/ Agriculture research Service, 1995. 125 p.
- BLASCO, A.; SORENSEN, D.; BIDANEL, J.P. Bayesian inference of genetic parameters and selection response for litter size components in pigs. **Genetics**, v. 149, p. 301-306, 1998.
- BLASCO, A. The bayesian controversy in animal breeding. **Journal of Animal Science**, v. 79, p. 2023-2046, 2001.
- DIAS, L.T.; ELFARO, L.; FRIES, L.A.; ALBUQUERQUE, L.G. Estimativas de parâmetros genéticos para perímetro escrotal e idade ao primeiro parto em animais da raça Nelore. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 37., 2000, Viçosa. **Anais...** Viçosa: SBZ, 2000. (CD-ROM).
- DIAS, D.S.O. **Análise genética de características de crescimento e de reprodução de animais da raça Nelore criados na região Centro-Oeste do Brasil.** 2001, 117 f. Tese (Doutorado) – Universidade Estadual Paulista, Jaboicabal, 2001.
- FARIA, C.U. **Inferência bayesiana no estudo genético quantitativo de características reprodutivas e de crescimento de bovinos da raça Nelore.** 2003. 73 f. Dissertação (Mestrado) – Universidade Federal de Goiás, Goiânia, 2003.

- GIANOLA, D.; FOULLEY, J.L. Variance estimation from integrated likelihood (VEIL). **Genetics Selection Evolution**, v. 22, p. 403-417, 1990.
- GUEDES, M.H.P.; MUNIZ, J.A.; SILVA, F.F.; AQUINO, L.H. Análise bayesiana da curva de crescimento de cordeiros da raça Santa Inês. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 57, p. 415-417, 2005.
- HOBERT, J.P.; CASELLA, G. The effect of improper priors on Gibbs sampling in hierarchical linear mixed models. **Journal of the American Statistical Association**, v. 91, p. 1461-1473, 1996.
- JOHNSON, N.L.; KOTZ, S. **Distributions in Statistics**. New York: John Wiley & Sons, 1972.
- KOOTS, K.R.; GIBSON, J.P.; SMITH, C. Analyses of published genetic parameter estimates for beef production traits: 1. Heritability. **Animal Breeding Abstracts**, v. 62, p. 309-338, 1994.
- LIN, E.C.; BERGER, P.J. Comparison of (co)variance component estimates in control populations of red flour beetle (*Tribolium castaneum*) using restricted maximum likelihood and Gibbs sampling. **Journal of Animal Breeding Genetics**, v. 118, p. 21-36, 2001.
- MAGNABOSCO, C.U. **Estimativas de parâmetros genéticos em características de crescimento de animais da raça Nelore usando os métodos de Máxima Verossimilhança Restrita e Amostragem Gibbs**. 1997, 89 f. Tese (Doutorado) – Universidade de São Paulo, Ribeirão Preto, 1997.
- MAGNABOSCO, C.U.; LÔBO, R.B.; FAMULA, T.R. Bayesian inference for genetic parameter estimation on growth traits for Nellore cattle in Brazil, using the Gibbs sampler. **Journal of Animal Breeding Genetics**, v. 117, p. 169-188, 2000.
- MARTINS, E.N.; LOPES, P.S.; SILVA, M.A.; TORRES, R.A.A. **Uso de modelos mistos na avaliação genética animal**. Viçosa: Editora UFV, 1997. 121 p. (Cadernos didáticos, 18).
- MELLO, S.P.; ALENCAR, M.M.; TORAL, F.L.B.; GIANLORENÇO, V.K. Estimativas de parâmetros genéticos para características de crescimento e produtividade em vacas da raça Canchim, utilizando-se inferência bayesiana. **Revista Brasileira de Zootecnia**, v. 35, p. 92-97, 2006.
- MERCADANTE, M.E.; LÔBO, R.B.; OLIVEIRA, H.N. Estimativas de (co)variâncias entre características de reprodução e de crescimento em fêmeas de um rebanho Nelore. **Revista Brasileira de Zootecnia**, v. 29, p. 997-1004, 2000.
- PEREIRA, E.; ELER, J.P.; COSTA, F.A.A.; FERRAZ, J.B.S. Análise genética da idade ao primeiro parto e do perímetro escrotal em bovinos da raça Nelore. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 53, p. 116-121, 2001.
- RESENDE, M.D.V. **Inferência bayesiana e simulação estocástica (Amostragem de Gibbs) na estimação de componentes de variância e valores genéticos em plantas perenes**. Colombo, PR: Embrapa Florestas, 2000. (Documentos 46).
- SILVA, J.A.V.; ELER, J.P.; FERRAZ, J.B.S.; OLIVEIRA, H.N. Análise genética da habilidade de permanência em fêmeas da raça Nelore. **Revista Brasileira de Zootecnia**, v. 32, p. 598-604, 2003.
- SORENSEN, D.A.; WANG, C.S.; JENSEN, J.; GIANOLA, D. Bayesian analysis of genetic change due to selection using Gibbs sampling. **Genetics Selection Evolution**, v. 26, p. 333-360, 1994.
- VAN KAAM, J.B.C.H.M. **Gibanal 2.9: analyzing Program for Markov Chain Monte Carlo Sequences**. Wageningen, The Netherlands: Dep. Anim. Sci., Wageningen Agricultural University, 1998.
- VAN TASSEL, C.P.; VAN VLECK, L.D. Multiple-trait Gibbs sampler for animal models: flexible programs for Bayesian and likelihood-based (co)variance component inference. **Journal Animal Science**, v. 74, p. 2586-2597, 1996.
- VAN TASSEL, C.P.; VAN VLECK, L.D.; GREGORY, K.E. Bayesian analysis of twinning and ovulation rates using a multiple trait threshold model and Gibbs sampling. **Journal Animal Science**, v. 76, p. 2048-2061, 1998.
- WANG, C.S.; RUTLEDGE, J.J.; GIANOLA, D. Marginal inferences about variance components in a mixed linear model using Gibbs sampling. **Genetics Selection Evolution**, v. 25, p. 41-62, 1993.
- WINTER, E.M.W.; ALMEIDA, M.I.M.; OLIVEIRA, E.G.; MARTINS, E.N.; NATEL, A.S.; SUREK, D. Aplicação do método bayesiano na estimação de correlações genéticas e fenotípicas de peso em codornas de corte em várias idades. **Revista Brasileira de Zootecnia**, v. 35, p. 1684-1690, 2006.