



**MODELO HIERÁRQUICO BAYESIANO
APLICADO NA AVALIAÇÃO GENÉTICA DE
CURVAS DE CRESCIMENTO DE BOVINOS
DE CORTE**

NATASCHA ALMEIDA MARQUES DA SILVA

2008

NATASCHA ALMEIDA MARQUES DA SILVA

**MODELO HIERÁRQUICO BAYESIANO APLICADO NA
AVALIAÇÃO GENÉTICA DE CURVAS DE CRESCIMENTO DE
BOVINOS DE CORTE**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de "Mestre".

Orientador

Prof. Dr. Renato Ribeiro de Lima

LAVRAS
MINAS GERAIS - BRASIL
2008

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Silva, Natascha Almeida Marques da.

Modelo hierárquico bayesiano aplicado na avaliação genética de curvas de crescimento de bovinos de corte / Natascha Almeida Marques da Silva. – Lavras: UFLA, 2008.

47 p. : il.

Dissertação (Mestrado) – Universidade Federal de Lavras, 2008.

Orientador: Renato Ribeiro de Lima.

Bibliografia.

1. Componentes de variância e covariância. 2. Metodologia hierárquica bayesiana. 3. Simulação. 4. Gado Nelore. I. Universidade Federal de Lavras. II. Título.

CDD – 519.542

NATASCHA ALMEIDA MARQUES DA SILVA

**MODELO HIERÁRQUICO BAYESIANO APLICADO NA AVALIAÇÃO
GENÉTICA DE CURVAS DE CRESCIMENTO DE BOVINOS DE CORTE**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Mestrado em Estatística e Experimentação Agropecuária, para a obtenção do título de “Mestre”.

APROVADA em 14 de fevereiro de 2008

Prof. Dr. Joel Augusto Muniz

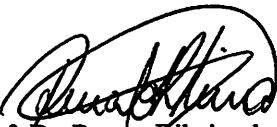
UFLA

Prof. Dr. Fabyano Fonseca e Silva

UFV

Prof. Dr. Rilke Tadeu Fonseca de Freitas

UFLA


Prof. Dr. Renato Ribeiro de Lima
UFLA

(Orientador)

LAVRAS

MINAS GERAIS-BRASIL

Aos meus pais Francisco e Eduarda

Dedico...

“Você não sabe o quanto eu caminhei.

Pra chegar até aqui...

Percorri milhas e milhas antes de dormir,

Eu nem cochilei

Os mais belos montes escalei,

Nas noites escuras de frio chorei....”

Trecho da música A Estrada do Cidade Negra

AGRADECIMENTOS

Agradeço a DEUS, por ter me guiado e protegido durante toda minha vida

Aos meus pais, Francisco e Eduarda, e aos meus irmãos, Thais e Pedro, pelo carinho, amor, alegria e incentivo, em todos os momentos da minha vida. Amo muito vocês!

Ao professor Renato, pela orientação, confiança, paciência, apoio, ensinamentos e, acima de tudo, pela amizade. Sou muito feliz por ter sido sua orientada.

Ao professor Luiz Henrique (Caveira), por ter sido mais que orientador em minha vida, por ter sido um grande amigo, pai e professor. Pelas maravilhosas reuniões em sua casa e, acima de tudo, por ter me ensinado com seus exemplos o que é ser profissional. Obrigada por tudo!

Ao professor Joel, por todos os ensinamentos, pela confiança, pela amizade e pelo apoio dado em todos esses anos.

Ao Fabyano, pela amizade, pelos ensinamentos e pelas risadas; por ter sido um grande apoio em todos os momentos desta jornada. Você e o Caveira foram fundamentais para eu me tornar o que sou hoje! Muito obrigada!

Ao Ricardo (Lirou), pela força e amizade constante e incansável, e pela ajuda fundamental na parte computacional.

Ao Luiz pelos conselhos e pelo carinho.

Aos meus queridos amigos do DEX, pelas festas, pelas risadas e pelos almoços de domingo. Vocês são minha família de Lavras!

Às meninas da graduação (ccs), pela amizade sincera e duradoura! Vocês são essenciais na minha vida! Amo vocês!

A minha linda família: tios, tias, primos, primas, avós e aos nossos bebês.

As minhas irmãs de coração, Cris, Lídia, Carol, Lílian e Lú, mesmo a distância, obrigada pela presença constante!

A Alininha, por ter sido minha amiga, irmã, filha, mãe e confidente. Você é muito especial! A Tenille que, mesmo com o pouco tempo de convivência, já faz parte da minha vida.

Aos professores e funcionários do Departamento de Ciências Exatas da UFLA.

À CAPES, pelo apoio com bolsas de estudo.

A todos que, de forma direta ou indireta, me auxiliaram neste trabalho.

SUMÁRIO

LISTA DE TABELAS.....	i
LISTA DE FIGURAS.....	ii
RESUMO.....	iii
ABSTRACT.....	iv
1 INTRODUÇÃO	1
2 REFERENCIAL TEÓRICO	3
2.1 Nelore	3
2.2 Modelos de crescimento.....	4
2.3 Inferencia Bayesiana.....	9
2.4 Métodos MCMC.....	11
2.4.1 Amostrador de Gibbs	12
2.4.2 Metropolis - Hastings.....	13
2.5 Estimação Utilizando a Inferência Bayesiana.....	16
2.6 Análise hierárquica Bayesiana.....	17
3 MATERIAL E MÉTODOS.....	18
3.1 Dados Simulados.....	18
3.2 Dados Observados.....	20
3.3 Processo de Estimação.....	21
4 Resultados e Discussão.....	26
4.2 Resultados obtidos para os dados simulados.....	26
4.3 Resultados obtidos para os dados reais.....	34
5 CONCLUSÕES.....	39
6 PROPOSTAS FUTURAS.....	39
7 REFERÊNCIAS BIBLIOGRÁFICAS.....	39

LISTA DE TABELAS

TABELA 1- Principais modelos não-lineares usados para o crescimento em bovinos.....	6
TABELA 2- Valores médios obtidos para as estimativas dos componentes de variância, herdabilidades, correlações genéticas e residuais das 18 repetições do processo de simulação.....	27
TABELA 3- Média, desvio padrão (DP), fator de dependência para o teste de Raftery e Lewis (FD), valor de z para o teste de Geweke (Z) e Erro de Monte Carlo (EMC) obtidas da densidade <i>a posteriori</i> das variâncias genéticas e ambientais, herdabilidades, correlações genéticas e ambientais para a repetição 8 do processo de simulação.....	29
TABELA 4- Média, desvio padrão (DP), intervalo de maior densidade <i>a posteriori</i> (HPD), fator de dependência para o teste de Raftery e Lewis (FD), valor de z para o teste de Geweke (Z) e Erro de Monte Carlo (EMC) obtidas da densidade <i>a posteriori</i> das variâncias genéticas e ambientais, herdabilidades e correlação genética para os parâmetros a e k para os dados de Nelore.....	35

LISTA DE FIGURAS

FIGURA 1- Valores paramétricos, estimados e limite inferior e superior do HPD para a herdabilidade do parâmetro a	32
FIGURA 2- Valores paramétricos, estimados e limite inferior e superior do HPD para a herdabilidade do parâmetro k	33
FIGURA 3 - Valores paramétricos, estimados e limite inferior e superior do HPD para a correlação genética entre os parâmetros a e k	33
FIGURA 4- Cadeias geradas pelo método MCMC e densidade a posteriori para $\sigma_g^2(a)$ e $\sigma_r^2(a)$ para a repetição 1.....	35
FIGURA 5- Cadeias geradas pelo método MCMC e densidade a posteriori para $\sigma_g^2(a)$ e $\sigma_r^2(a)$ para a repetição 8.....	36
FIGURA 6- Cadeias geradas pelo método MCMC e densidade a posteriori para $\sigma_g^2(a)$ e $\sigma_r^2(a)$ para os dados de Nelore.....	40
FIGURA 7- Cadeias geradas pelo método MCMC e densidade a posteriori para $\sigma_g^2(k)$ e $\sigma_r^2(k)$ e para os dados de Nelore.....	41

RESUMO

SILVA, Natascha Almeida Marques da. **Modelo hierárquico bayesiano aplicado na avaliação genética de curvas de crescimento**. 2008. 47 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, MG.*

A estimação dos componentes de (co)variância dos parâmetros de modelos de crescimento pode ser feita por vários métodos. A metodologia bayesiana se apresenta como uma forma alternativa de estimação. Foi realizado um estudo por meio de dados simulados e de dados reais de animais do gado Nelore, para a estimação dos componentes de (co)variância dos parâmetros do modelo de Von Bertalanffy, por meio da metodologia hierárquica bayesiana. A partir dos componentes estimados foram encontradas as herdabilidades para cada parâmetro do modelo e as correlações genéticas e ambientais entre esses parâmetros. As distribuições marginais a posteriori dos parâmetros a , R , μ , u , G e σ_c^2 foram obtidas por meio do algoritmo Gibbs Sampler e as dos parâmetros b e k por meio do algoritmo Metropolis-Hastings. A metodologia se mostrou eficiente, proporcionando estimativas para os parâmetros próximas aos valores simulados. Os parâmetros a e k dos dados reais apresentaram valores de herdabilidades compatíveis com a realidade, indicando que esses parâmetros poderiam ser usados para fins de seleção.

Palavras-chave: Componentes de variância e covariância, metodologia hierárquica bayesiana, simulação e gado Nelore.

* Comitê Orientador: Prof. Renato Ribeiro de Lima- UFLA (orientador), Prof. Joel Augusto Muniz - UFLA e Prof. Luiz Henrique de Aquino - UFLA.

ABSTRACT

SILVA, Natascha Almeida Marques da. **Bayesian hierarchical model applied to genetic evaluation of growth curves of beef cattle.** 2008. 47 p. Dissertation (Master Program in Statistics and Agricultural Experimentation) - Federal University of Lavras, Lavras, Minas Gerais, Brazil.*

The estimation of the (co)variance components for the parameters of the growth models can be evaluated by many methods. The Bayesian approach report as an alternative method of the estimation. We performed a study by simulation structure with application on to Nelore cattle data for estimation of the (co)variance components for the parameters of Von Bertalanffy growth curve, by using a Bayesian hierarchical model. From the estimated components were found the heritabilities for each parameter and genetics and environmentals correlations between these parameters. The samples of posterior marginal distributions for the parameters a , R , μ , u , G and σ_e^2 were obtained via Gibbs Sampler algorithm and for the parameters b e k by the Metropolis-Hastings algorithm. The efficiency of the Bayesian Inference methodology was verified since estimated parameters were quite close the simulated ones. The parameters a and k from real data showed heritabilities compatibles with the reality indicating they could be used in selection programs.

Key words: (co)variance components, Bayesian hierarchical methodology, simulation and Nelore cattle

*Guidance Committee: Prof. Renato Ribeiro de Lima- UFLA (adviser), Prof. Joel Augusto Muniz – UFLA and Prof. Luiz Henrique de Aquino – UFLA.

1 Introdução

O aumento do peso vivo ao longo da vida dos animais é um fenômeno complexo e dependente dos efeitos ambientais e genéticos que atuam ao longo do tempo. De acordo com Tedeschi et al. (2000), o conhecimento e o controle do crescimento e do desenvolvimento dos ruminantes são tópicos de bastante interesse para os criadores e pesquisadores. O domínio desses fatores permite que o manejo nutricional dos animais possa ser conduzido eficientemente, além de permitir que programas de seleção animal sejam elaborados para as características de crescimento dos animais.

O crescimento apresenta relação direta com a quantidade e a qualidade da carne. Assim, os criadores de gado de corte estão cada vez mais conscientes da importância da avaliação do crescimento. Uma das formas de se avaliar o crescimento animal é por meio de curvas de crescimento. Essas curvas são obtidas a partir de um conjunto de pesagens tomadas em um mesmo indivíduo, em intervalos regulares ou não, desde o nascimento até a idade adulta.

Embora curvas de crescimento seja um tema antigo, as dificuldades dos cálculos inviabilizavam sua aplicação. Atualmente, com os avanços computacionais, métodos mais rápidos e eficientes foram desenvolvidos, facilitando o estudo das curvas de crescimento. A seleção para alterar a forma da curva de crescimento dos animais, aumentando a taxa de crescimento até ao peso de abate sem alterar o tamanho adulto, tem sido preconizada como uma alternativa para melhorar a eficiência de produção em gado de corte.

Segundo Oliveira (2000), modelos matemáticos não lineares, desenvolvidos empiricamente, têm se mostrado adequados para descrever curvas de crescimento. Em bovinos, os cinco mais utilizados são: Brody, Von Bertalanffy, Richards, Logístico e Gompertz. Em geral, ajustam-se os modelos não-lineares a um conjunto de dados (peso e idade, por exemplo) com a

finalidade de prever a variável dependente (peso) para valores da variável independente (idades). Estes modelos fornecem parâmetros, cujas estimativas são importantes para descrever o crescimento animal.

Além dos parâmetros que são estimados diretamente do modelo e que apresentam interpretação biológica, podem ser obtidos também os parâmetros genéticos referentes a eles, como as herdabilidades e correlações, as quais podem ser utilizadas em programas de melhoramento animal.

Vários autores, dentre eles DeNise e Brinks (1985), Elias (1998), Oliveira et al. (2000) e Silva et al. (2002), registraram que a herdabilidade para as estimativas dos parâmetros do modelo é de moderada a alta, indicando que estas características podem responder bem a programas de seleção. Porém, quando se dispõem de animais sujeitos a diferentes efeitos ambientais ou genéticos, se torna necessário que as estimativas dos parâmetros sejam corrigidas para estes efeitos. Esta correção é extremamente útil quando o objetivo é alterar a forma da curva de crescimento para programas de melhoramento animal que visam ao aumento da eficiência do crescimento.

As estimativas dos parâmetros dos modelos de crescimento em programas de seleção, normalmente, são obtidas por meio de uma metodologia frequentista, que considera duas fases distintas: a primeira corresponde ao ajuste individual dos modelos não-lineares, de forma a obter um conjunto de estimativas de parâmetros para cada animal e a segunda corresponde à caracterização dessas estimativas como variáveis dependentes em um modelo misto contendo efeitos fixos e aleatórios, de forma a se obter os componentes de (co)variância necessários ao cálculo das herdabilidades e das correlações.

Um fato importante decorrente da aplicação dessa metodologia é que os fatores genéticos e ambientais não são considerados no processo de estimação individual, ou seja, as estimativas dos parâmetros não são corrigidas simultaneamente para estes fatores, o que implica em perdas de informações.

Essas perdas prejudicam o processo de seleção de animais geneticamente superiores. Uma perda importante é aquela relativa aos animais para os quais não se conseguem suas curvas de crescimento estimadas, geralmente por problemas de convergência. Esses animais, mesmo não apresentando convergência, serão considerados na segunda etapa do processo de estimação, devido ao uso de modelos mistos, mas perde-se em acurácia.

Uma possível solução para este problema é a utilização da Inferência Bayesiana considerando um modelo hierárquico em vários estágios. Essa metodologia contempla a estimação simultânea dos parâmetros e dos fatores genéticos e ambientais que atuam sobre eles, havendo assim uma correção instantânea para esses fatores, devida às atualizações relativas ao processo iterativo (Silva et al., 2008a).

O objetivo da realização deste trabalho foi utilizar a inferência bayesiana para estimar os componentes de (co)variância dos parâmetros do modelo de Von Bertalanffy e, a partir desses componentes, obter as herdabilidades e as correlações genéticas e ambientais para cada um desses parâmetros. Essa estimação foi feita considerando-se dados simulados e dados observados de animais da raça Nelore.

2 -Referencial teórico

2.1 Nelore

A raça Nelore, segundo Santos (1999), correspondente ao ongole da Índia, predomina atualmente no cenário pecuário brasileiro. O primeiro registro de Nelore no Brasil aconteceu em 1868, quando um navio, que se dirigia à Inglaterra, ancorou em Salvador com um casal de reprodutores a bordo, os quais acabaram sendo comercializados. Aos poucos a raça foi se expandindo e, hoje,

está presente em todos os confinamentos do país, sendo a principal raça utilizada para cruzamentos industriais.

O Nelore é um animal adaptado às condições dos trópicos, enfrentando os períodos de seca que duram até seis meses, quando se alimenta de pasto de baixa qualidade, além dos ecto e endoparasitos.

Na região Centro-Oeste, onde se concentra a maior parte do gado Nelore, a produtividade é intensa. As fêmeas parem com extrema facilidade, criam os seus bezerros e continuam produzindo até os 20 anos de idade. É uma raça rústica, fértil, prolífera, de longa vida reprodutiva e resistente às doenças comuns dessa região.

O Nelore tem pêlos curtos, finos e lisos, que auxiliam na eliminação do calor. A pelagem branco-cinza e a pele preta apresentam um conjunto de propriedades físicas capazes de refletir, absorver, irradiar e filtrar as diversas radiações solares dos trópicos.

O ganho de peso desta raça é uma característica de alta herdabilidade ($h^2 = 40\%$) e os pesquisadores já constataram a potencialidade do Nelore para ganhar peso e acrescentar mais carne nos seus descendentes (Giannotti et al., 2005). Os criadores têm buscado então através da seleção, animais com maior ganho de peso em menor intervalo de tempo e assim obter maior lucro na atividade.

Além do ganho de peso, várias outras características relacionadas ao peso dos animais são objeto de estudo, como: peso ao nascimento, peso à desmama, peso ao ano, ao sobreano, etc. Porém, é de grande importância estudá-las conjuntamente e isso se torna possível com o uso de modelos de crescimento (Toral et al., 2007).

2.2 Modelos de crescimento

Segundo Fitzhugh (1976), uma curva sigmóide descrevendo uma seqüência de medidas de tamanho, freqüentemente peso corporal, sugere o termo “curva de crescimento” ou “curvas de peso-idade”. O mesmo autor destaca que os principais objetivos no ajuste de curvas de crescimento são descritivos, de modo que a informação contida numa seqüência de pontos peso-idade é reduzida em relativamente poucos parâmetros que tenham interpretação biológica.

Os parâmetros da curva de crescimento são utilizados para prever taxas de crescimento, necessidades alimentares, peso à maturidade e graus de maturidade, também servindo como critérios de seleção para programas de melhoramento animal. Nesse sentido, as curvas de crescimento refletem a relação entre a idade do animal e o seu impulso de crescimento e maturidade, sendo importantes para pesquisas e recomendações sobre eficiência de produção em programas de melhoramento, contribuindo, dessa maneira, para aumentar o lucro do produtor (Mendes, 2007).

As funções de crescimento são capazes de condensar informações de dados de peso - idade de todo o período de vida de um animal em um conjunto de parâmetros biologicamente interpretáveis (Laird e Howard, 1967). Este fato, associado às características produtivas dos animais, pode ser uma alternativa para programas de seleção, visando à precocidade com maior peso e melhor qualidade de carcaça (Souza e Bianchini Sobrinho, 1994).

O interesse pelas curvas de crescimento aumentou muito nos últimos anos, principalmente pelo desenvolvimento de novas técnicas computacionais que permitem maior rapidez e precisão das análises e também devido à maior necessidade de se prever informações sobre um rebanho (Silva et al., 2001).

Uma das formas de se descrever as curvas de crescimento é por meio de modelos de regressão não linear. Os principais, de acordo com Mendes (2007), estão listados na Tabela 1:

TABELA 1 Principais modelos não lineares usados para o crescimento em bovinos.

Função	Equação	Nº de parâmetros
Brody	$a(1 - b \exp(-kt))$	3
Von Bertalanffy	$a(1 - b \exp(-kt))^3$	3
Logística	$a(1 + b \exp(-kt))^{-1}$	3
Gompertz	$a \exp(-b \exp(-kt))$	3
Richards	$a(1 - b \exp(-kt))^m$	4

De acordo com os modelos apresentados na Tabela 1, o parâmetro a é o valor assintótico da função quando t (tempo, idade) tende ao infinito. Genericamente, este valor representa o peso adulto do animal, independente de problemas de flutuações ou, mesmo, de efeitos climáticos.

O parâmetro b , ou constante de integração, não possui interpretação biológica e é utilizado para adequar o valor inicial do peso vivo, nesse caso, peso ao nascer da função utilizada.

O parâmetro k é uma função entre a taxa média de crescimento e o peso adulto do animal, geralmente sendo referido como taxa de maturidade. A taxa de maturidade pode ser entendida como um indicador da velocidade com que o animal se aproxima de seu tamanho adulto. Portanto, quanto maior este valor, mais precoce é o animal, em termos de crescimento.

O parâmetro m é o que dá forma à curva de crescimento e, conseqüentemente, determina seu ponto de inflexão, que representa o momento em que se encerra a aceleração do crescimento e se inicia a fase de auto-inibição até chegar ao peso adulto. Assume os seguintes valores para os modelos de crescimento: $m = 1$ (Brody); $m = 3$ (Von Bertalanffy); $m = \infty$ (Gompertz) e $m = -1$ (Logístico).

Forni (2007) ressalta que, embora as letras utilizadas para designar cada um dos parâmetros coincidam nos diferentes modelos, existem variações quanto às suas interpretações. Em alguns casos, a interpretação dos parâmetros é direta, mas, em outros, é necessário recorrer à dedução da equação para uma compreensão exata de sua relação com os fenômenos fisiológicos subjacentes. As comparações de estimativas de parâmetros obtidas em diferentes modelos, ainda que sejam parâmetros designados pela mesma letra, podem estar equivocadas. Apesar disso, associações biológicas podem ser feitas.

De acordo com Santoro et al. (2005), o benefício do uso de modelos não-lineares no melhoramento genético animal poderia ser observado na estimação dos parâmetros das curvas, proporcionando a identificação dos animais mais adequados para determinados fins de seleção. Forni (2007) afirma que o interesse no parâmetro a advém da impossibilidade de medir o peso de todos os animais na idade adulta, pois o peso ao abate é atingido antes da maturidade. Em relação ao parâmetro k , a maturação precoce é uma característica desejável na produção de bovinos de corte, uma vez que representa diminuição no tempo que o animal necessita para atingir as características desejáveis para o abate, melhorando assim a taxa de retorno da atividade pecuária.

Dentre os modelos apresentados na Tabela 1, pode-se considerar, em relação aos aspectos matemáticos, o modelo Brody como sendo o mais simples, uma vez que ele não apresenta ponto de inflexão, ou seja, o crescimento ocorre de forma constante. Porém, esta não é a realidade do crescimento de bovinos de corte. Por outro lado, o modelo Richards apresenta-se como o mais complexo, pois o parâmetro que define seu ponto de inflexão (m) é variável e este fato, muitas vezes, implica em uma grande dificuldade de obtenção das estimativas (Silva et al., 2001). De forma geral, o modelo de Bertalanffy caracteriza-se como

sendo eficiente, pois apresenta formulação matemática relativamente simples, inclusive com a definição de um ponto de inflexão, cujo valor é fixo.

Segundo Forni (2007), Von Bertalanffy desenvolveu sua função baseando-se na suposição de que o crescimento é a diferença entre as taxas de anabolismo e catabolismo dos tecidos, que teoricamente poderiam ser conhecidos mediante experimentos de laboratório.

Carrijo e Duarte (1999), em suas revisões, destacam que a função de Von Bertalanffy foi a mais apropriada para descrever o crescimento e o desenvolvimento de animais da raça Nelore. Os critérios citados para destacá-la como melhor foram: qualidade no ajuste, ausência de valores irreais de peso adulto e facilidade computacional.

Dos parâmetros dos modelos de crescimento, podem ser obtidas as estimativas de seus componentes de (co)variância. A partir dessas estimativas obtêm-se as correlações e herdabilidades para cada parâmetro do modelo. A herdabilidade de um caráter quantitativo em uma população é o parâmetro genético de maior importância, já que determina a estratégia a ser usada no melhoramento deste caráter. A herdabilidade é o quociente de variância genética aditiva sobre a soma da variância genética aditiva com a variância residual. Assim, ao diminuir a variância residual, ou seja, ter um melhor controle das condições de meio, pode-se notar que determinadas características são mais herdáveis (Elias, 1998).

Segundo Faria et al. (2006), em suas revisões, estimativas acuradas dos componentes de (co)variância têm sido altamente desejáveis na produção animal, uma vez que as diferenças entre o valor verdadeiro e o valor estimado aumentam o erro de predição dos efeitos aleatórios (valores genéticos).

Entretanto, Forni (2007) afirma que existem poucos trabalhos a respeito dos componentes de (co)variâncias de parâmetros de curvas de crescimento de zebuínos. Dos que estão disponíveis no Brasil, há uma grande variação nas

estimativas feitas, provavelmente porque muitos deles foram realizados com pequeno número de dados e em muitos foram feitas extrapolações a partir de pesagens obtidas até os 24 meses de idade. Santoro et al. (2005) também concordam com o pouco número de trabalhos existente e afirmam que a maioria dos autores estudou somente o ajuste das curvas, disponibilizando apenas os valores médios dos parâmetros para as amostras analisadas, não relatando a respeito dos ajustes individuais.

Carrijo e Duarte (1999) relatam que a seleção para o crescimento baseada somente nos valores fenotípicos dos parâmetros das curvas de crescimento pode não levar a uma resposta satisfatória. Assim, o conhecimento do comportamento genético de cada parâmetro do modelo é extremamente necessário para a seleção de animais em um rebanho.

A função de crescimento de Brody foi ajustada por Silva et al. (2002) a dados de crescimento de novilhos Nelore para estimar a herdabilidade dos parâmetros a (0,28) e k (0,36). A partir das estimativas obtidas, estes autores sugerem que estes parâmetros podem ser utilizados no estabelecimento de programas de seleção.

Garnero et al. (2005), estudaram o crescimento de fêmeas Nelore pelos modelos de Brody, Gompertz, Logístico e Von Bertalanffy e concluíram que os coeficientes de herdabilidade foram de elevada magnitude para o parâmetro a (0,39-0,42) e também para o parâmetro k (0,39-0,44), em todos os modelos. Assim, tem-se um indicativo de que estes poderiam ser usados como instrumentos para alterar a forma da curva de crescimento destes animais, por seleção.

Talhari et al. (2003) a partir dos parâmetros do modelo de Von Bertalanffy, estimaram os parâmetros genéticos de fêmeas da raça Canchim pelo método da máxima verossimilhança restrita e também concluíram que a seleção para características ligadas ao crescimento resulta em progresso genético.

Em relação às correlações, a de maior importância é a correlação genética, que relaciona os parâmetros a e k , que são os que possuem interpretação biológica. Em geral, a correlação entre estes parâmetros é alta e negativa. Assim, variações genéticas que estariam ligadas a maiores pesos adultos também estariam associadas a taxas de maturação mais lentas (Forni, 2007).

Vários são os procedimentos de estimação dos componentes de (co)variância já propostos para aplicação no melhoramento animal. Faria et al. (2007) relatam que, entre eles, estão métodos freqüentistas, como as análises de variância (ANAVA), os métodos de máxima verossimilhança (MV), de máxima verossimilhança restrita (MVR) e os métodos bayesianos. A metodologia bayesiana tem sido proposta como uma alternativa à abordagem freqüentista para a estimação dos componentes de (co)variância.

2.3 Inferência bayesiana

A escola bayesiana foi fundada por Laplace, por meio de vários trabalhos publicados de 1774 a 1812 e, durante o século XIX, desempenhou papel preponderante na inferência científica (Stigler, 1986). Mas, por necessitar de resoluções de funções matemáticas inviáveis de serem obtidas analiticamente, a metodologia bayesiana ficou resguardada durante um grande período.

Por volta de meados do século XX, mais precisamente na década de 1960, a análise bayesiana ressurgiu em alguns trabalhos teóricos, como o de Jeffreys (1961). Mas somente em 1990, com o trabalho de Gelfand e Smith (1990), que exploraram um recurso de simulação dinâmica, o algoritmo Gibbs Sampler, é que o problema das integrações foi solucionado de maneira alternativa. Em melhoramento genético animal, o método bayesiano foi

introduzido por Daniel Gianola, desenvolvendo trabalhos em praticamente todas as áreas do melhoramento animal, na década de 1990 (Faria et al., 2007).

No contexto da inferência bayesiana, todos os parâmetros do modelo, sejam fixos ou aleatórios, são vistos como variáveis aleatórias, conforme o conceito subjetivo de probabilidade. Levando em conta a incerteza existente sobre todos os parâmetros simultaneamente, a obtenção de intervalos de credibilidade é imediata e é, geralmente, mais precisa que os intervalos obtidos pela metodologia frequentista que, em geral, utiliza variâncias assintóticas e aproximadas (O'Hagan, 1994). Além disso, a utilização dessa metodologia requer um número menor de observações, pois os conceitos probabilísticos envolvidos diminuem a dependência do ajuste do modelo em relação ao número de observações utilizadas, uma vez que o conceito de graus de liberdade não é utilizado (Silva et al., 2005 e Silva et al., 2006).

A metodologia bayesiana utiliza as informações *a priori* sobre os parâmetros e a informação dos dados amostrais. A partir dessas informações, calcula a densidade *a posteriori* dos parâmetros a serem estimados. As informações *a priori* expressam o conhecimento que o pesquisador tem sobre os parâmetros a serem estimados e são dadas pela densidade de probabilidade $P(\theta)$, em que θ é o vetor de parâmetros a serem estimados. Quando em determinado estudo o pesquisador tem pouca ou nenhuma informação para incorporar *a priori*, considera-se uma *priori* não-informativa, por exemplo, a *priori* de Jeffreys (Jeffreys, 1961). Os dados $Y = \{y_1, y_2, \dots, y_n\}$, representados por uma amostra aleatória de uma população com densidade f , são considerados na análise bayesiana por meio da função de verossimilhança $L(Y|\theta)$, que é a densidade conjunta destes dados (Silva et al., 2008b).

Assim, por meio do Teorema de Bayes, a partir da distribuição *a priori* e da função de verossimilhança é possível obter a distribuição de densidade *a*

posteriori $P(\theta | Y)$, de forma que qualquer conclusão a respeito dos parâmetros é obtida a partir desta distribuição. Portanto, temos que:

$$P(\theta | Y) \propto L(Y | \theta) P(\theta)$$

ou seja: *Posteriori* \propto *Verossimilhança* \times *Priori*, em que \propto representa proporcionalidade.

Segundo Broemiling (1989), pode-se pensar no Teorema de Bayes como um mecanismo de atualização da opinião do pesquisador sobre θ . Esse teorema constitui a base da inferência bayesiana, pois toda prática inferencial é realizada a partir da distribuição *a posteriori* obtida.

Segundo Rosa (1998), para se inferir em relação a qualquer elemento de θ , a distribuição *a posteriori* conjunta dos parâmetros, $P(\theta | Y)$, deve ser integrada em relação a todos os outros elementos que a constituem, obtendo-se a distribuição marginal de cada elemento.

A integração da distribuição conjunta *a posteriori* para a obtenção das marginais, geralmente, não é analítica, sendo muito difícil e trabalhosa de ser solucionada por métodos numéricos. Utilizam-se, então, algoritmos especializados que utilizam os métodos denominados MCMC (Markov Chain - Monte Carlo). Para a utilização desses algoritmos, é necessário que se obtenha, a partir da distribuição *a posteriori*, um conjunto de distribuições chamadas de distribuições condicionais completas (Silva et al., 2008b).

2.4 Métodos MCMC

A introdução de métodos de Monte Carlo baseados em cadeias (seqüências) de Markov é denominada MCMC. Os métodos MCMC utilizam a simulação Monte Carlo para a geração de valores para os parâmetros a partir de distribuições de probabilidades. Fazem uso também da teoria das Cadeias de

Markov, que considera um sistema de dependência entre observações subsequentes geradas na simulação. O Gibbs Sampler e o Metropolis-Hastings são algoritmos que utilizam esses métodos.

2.4.1 Gibbs Sampler (Amostrador de Gibbs)

O Gibbs Sampler é um algoritmo que gera amostras aleatórias de uma distribuição marginal, sem que se conheça sua densidade (Gamerman & Lopes (2006). Em modelos mais complexos, raramente se conseguem obter amostras diretamente das distribuições *a posteriori*.

É essencialmente um esquema iterativo de amostragem, cujo núcleo de transição é formado pelas distribuições condicionais completas. À medida que o número q de iterações aumenta, a seqüência de valores gerados se aproxima da distribuição de equilíbrio, ou seja, da densidade marginal desejada para cada parâmetro (Silva, 2006).

Suponhamos que o vetor de parâmetros θ seja dividido em k subvetores, $(\theta_1, \theta_2, \dots, \theta_k)'$, e que as distribuições condicionais de cada parâmetro θ_i , $i = 1, 2, \dots, k$, dado todos os outros, sejam representadas por distribuições condicionais conhecidas. Estas distribuições são denotadas por: $f_1(\theta_1 | \theta_2, \dots, \theta_k, Y)$, $f_2(\theta_2 | \theta_1, \theta_3, \dots, \theta_k, Y)$, ..., $f_k(\theta_k | \theta_1, \dots, \theta_{k-1}, Y)$.

Segundo Gamerman & Lopes (2006), o Amostrador de Gibbs pode ser descrito nos seguintes passos:

(1) definem-se os valores iniciais $\theta^0 = (\theta_1^0, \theta_2^0, \dots, \theta_k^0)'$ para os parâmetros;

(2) Amostra-se-se iterativamente

$$\theta_i^{(l)} \text{ de } f_i(\theta_i | \theta_2^{(l)}, \theta_3^{(l)}, \dots, \theta_k^{(l)}, Y)$$

$$\theta_2^{(1)} \text{ de } f_2(\theta_2 | \theta_1^{(1)}, \theta_3^{(0)}, \dots, \theta_k^{(0)}, \mathbf{Y})$$

⋮

$$\theta_k^{(1)} \text{ de } f_k(\theta_k | \theta_1^{(1)}, \theta_2^{(1)}, \dots, \theta_{k-1}^{(1)}, \mathbf{Y})$$

obtendo-se, na primeira iteração, $\theta^1 = (\theta_1^{(1)}, \dots, \theta_k^{(1)})'$. Cada passo desses é chamado de iteração;

- (3) repete-se o passo (2) q vezes, ou seja, após um grande número de iterações (q iterações), obtém-se $\theta^q = (\theta_1^{(q)}, \dots, \theta_k^{(q)})'$.

O conjunto dos q valores amostrados representam amostras da distribuição marginal *a posteriori* de θ . A partir desta distribuição obtém-se as estimativas pontuais (médias, medianas e modas) e estimativas por região (intervalos de credibilidade e HPD) para os parâmetros amostrados. O HPD é denominado intervalo de credibilidade de máxima densidade *a posteriori*. Quanto menor for o tamanho do intervalo, mais concentrada é a distribuição do parâmetro e a amplitude do intervalo informa sobre a dispersão de θ .

2.4.2 Metropolis-Hastings

O algoritmo Metropolis-Hastings é utilizado para a obtenção da distribuição marginal a posteriori quando o amostrador de Gibbs não se mostra eficiente, ou seja, para parâmetros (θ) cuja distribuição condicional não se caracteriza como uma distribuição de probabilidades conhecida. Neste caso, geram-se valores do parâmetro (β) a partir de uma distribuição proposta, cuja densidade seja $q(\theta, \beta)$, e esse é aceito ou não com uma certa probabilidade de aceitação.

Segundo Gamerman e Lopes (2006), pode ser descrito nos seguintes passos:

(1) inicializar com um valor arbitrário $\theta^{(0)}$, $j = 0$;

(2) Geram-se β de $q(\theta^{(j)}, \cdot)$ e u de uma uniforme $(0,1)$;

(3) Calcula-se $p = \min \left\{ 1, \frac{P(\beta)q(\theta^{(j)}, \beta)}{P(\theta^{(j)})q(\beta, \theta^{(j)})} \right\}$;

(4) se $u \leq p$, faz-se $\theta^{(j+1)} = \beta$, caso contrário $\theta^{(j+1)} = \theta^{(j)}$;

(5) repetem-se os passos (2) a (4), até que a distribuição de equilíbrio tenha sido obtida.

À medida que o número de iterações aumenta, $q \rightarrow \infty$, a seqüência se aproxima de sua condição de equilíbrio, ou seja, ela atinge a característica de estacionariedade com um comportamento assintótico (Nogueira et al., 2004).

Para a implementação dos algoritmos Gibbs Sampler e Metropolis-Hastings, são necessários cuidados especiais. Assim, as seguintes definições devem ser feitas: convergência, período de descarte amostral (*burn-in*), intervalo de amostragem (*thin*) e tamanho de cadeia amostral.

a) Convergência

Como os algoritmos Gibbs Sampler e Metropolis-Hastings são processos iterativos, deve-se constatar a convergência dos mesmos. Teoricamente, o processo iterativo teria fim apenas quando o número de iterações tendesse ao infinito. Na prática admite-se convergência quando a série alcança um estado de estacionariedade, o que significa que as condicionais completas estão suficientemente próximas das distribuições marginais. Para a verificação formal da convergência existe alguns critérios descritos na literatura, como os critérios

de Geweke (1992) e Raftery Lewis (1992), dentre outros. Estes critérios encontram - se muito bem detalhados em Nogueira et al. (2004).

b) Período de descarte amostral

O período de descarte amostral (*burn-in*) é o número de iterações que devem ser descartadas da parte inicial da cadeia. A justificativa para o *burn-in* se baseia no fato de que à medida que o número de iterações aumenta, a cadeia se torna independente dos valores iniciais e converge para a distribuição de equilíbrio.

c) Intervalo de amostragem

Para se estimar parâmetros de uma distribuição, uma amostra de observações independentes se torna necessária. Amostras consecutivas, obtidas das cadeias do Gibbs Sampler e Metropolis-Hastings, podem ser altamente correlacionadas. Dessa forma, para se obter uma amostra independente, ou não correlacionada, é necessário que se eliminem resultados de iterações intermediárias da cadeia amostral. O intervalo de retirada da amostra é denominado intervalo de utilização amostral (*thin*) e deve ser suficiente para que as amostras utilizadas não sejam correlacionadas.

d) Tamanho de cadeia amostral

Os algoritmos Gibbs Sampler e Metropolis-Hastings produzem cadeias de amostras com tantas iterações quantas forem especificadas. Este aspecto é muito importante, porque está intimamente relacionado com a convergência das distribuições marginais *a posteriori* dos parâmetros. Se o número de iterações

definido pelo usuário for adequado, as médias posteriores das amostras serão estimativas válidas das distribuições marginais *a posteriori* dos parâmetros.

2.5 Estimação usando a inferência bayesiana

A diferença formal entre a inferência bayesiana e a freqüentista é que, para a inferência bayesiana, o parâmetro θ é variável aleatória, possuindo, então, uma distribuição de probabilidade. E, para a inferência freqüentista, os parâmetros são valores fixos, não sendo possível atribuir a eles uma distribuição de probabilidade (Toral et al., 2007)

Faria et al. (2006) compararam a estimação dos componentes de (co)variância pelo método de máxima verossimilhança restrita e pelo método bayesiano e concluíram que este último possui vantagem em relação ao primeiro devido à obtenção das distribuições marginais que oferecem mais informações a respeito dos parâmetros.

De acordo com Faria et al. (2007), com o uso dos métodos freqüentistas, como é o caso do método de máxima verossimilhança restrita, obtêm-se somente estimativas pontuais da variância genética. Assim, diante de uma simples medida de erro, que só teria sentido em amostras muito grandes, e de dados distribuídos de forma normal em análises em que não há o conhecimento das médias e variâncias, haveria de se utilizar as estimativas da máxima verossimilhança como se fossem os parâmetros verdadeiros, ignorando o seu erro. As análises bayesianas, porém, permitem calcular a densidade marginal *a posteriori* do parâmetro de interesse e, com tal distribuição, quantifica-se exatamente o erro, apresentando maior precisão nas estimativas obtidas.

Toral et al. (2007) descrevem que os componentes de (co)variância estimados via máxima verossimilhança restrita e inferência bayesiana podem ser semelhantes, especialmente para características com alta herdabilidade. Todavia,

as estimativas de erro podem ser menores para os componentes obtidos por inferência bayesiana, em virtude da influência da distribuição *a priori* sobre os componentes de (co)variância *a posteriori*. Porém, essas diferenças diminuem quando o número de animais envolvidos nas análises aumenta.

Segundo Van Tassell et al. (1995), a utilização da inferência bayesiana permite obter, além das estimativas, o intervalo de credibilidade para a distribuição *a posteriori* dos componentes de (co)variância. Além disso, ela pode ser mais vantajosa que o método de máxima verossimilhança restrita quando o arquivo de dados é muito grande e os modelos são bastante complexos.

2.6 Análise hierárquica bayesiana

Um modelo hierárquico é aquele que especifica estágios, com cada estágio sendo construído sobre outro. A vantagem da construção de um modelo em estágios é que cada um pode ser relativamente simples e de fácil entendimento, enquanto um único modelo pode ser mais complexo (Searle, Casella, McCulloch, 1992). Segundo esses autores, a estimação bayesiana está fortemente ligada à definição de modelos hierárquicos.

Varona et al. (1997) sugeriram, pela primeira vez na literatura, a análise hierárquica bayesiana visando estudar a influência de efeitos genéticos e ambientais sobre parâmetros de qualquer função de produção (índices de crescimento, curvas de crescimento, curvas de lactação, etc.) e descreveram o algoritmo de Gibbs sampler para a implementação desses procedimentos. Estes autores mostraram a aplicação da metodologia em uma função de crescimento linear com dois parâmetros, ajustada a dados de peso-idade de bovinos de corte da raça Pirenaica e observaram que a metodologia proporcionou redução no número de curvas atípicas. Além disso, concluíram os autores que a aplicação

desses procedimentos possibilita alterar a forma de curvas de crescimento de funções de produção por seleção genética.

Varona et al. (1999) por meio de simulação de dados, compararam a eficiência da análise hierárquica bayesiana em relação à análise realizada em dois estágios independentes. Nesta última análise, ajustam-se, primeiramente, curvas individuais e, posteriormente, identificam-se efeitos ambientais e genéticos sobre os parâmetros do modelo de crescimento. Os resultados mostraram que, nesta última, houve perdas de informações por não considerar diretamente efeitos fixos e aleatórios no processo de obtenção das curvas.

Em recentes estudos envolvendo ajuste de modelos não-lineares a dados reais de produção animal (Blasco et al., 2003; Chang, et al., 2001; Silva et al., 2005; Guedes et al., 2005, Forni, 2007 e Muniz et al., 2007), a inferência bayesiana foi utilizada com sucesso, pois reduziu o número de estimativas viesadas, mesmo utilizando poucas informações; permitiu a estimação de variações existentes entre os animais simultaneamente com a estimação dos parâmetros do modelo de crescimento e apresentou estimação por intervalo mais precisa em relação à obtida pela metodologia frequentista.

3 Material e Métodos

3.1 Dados simulados

A simulação foi conduzida através de populações geradas em 2 etapas distintas, segundo procedimentos apresentados por Varona et al. (1999).

Na primeira, os parâmetros da função de Von Bertalanffy foram gerados por meio de uma distribuição normal multivariada cujos parâmetros são definidos pelo modelo animal e tem a seguinte estrutura:

$$\begin{pmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{k} \end{pmatrix} \sim N \begin{pmatrix} \mu_a + \mathbf{Zu}_a \\ \mu_b + \mathbf{Zu}_b \\ \mu_k + \mathbf{Zu}_k \end{pmatrix}, \mathbf{R} \otimes \mathbf{I}$$

em que :

\mathbf{a} , \mathbf{b} , e \mathbf{k} são os vetores de parâmetros da função de Von Bertalanffy, com:

$\mathbf{a} = [a_1 \ a_2 \ \dots \ a_N]'$, $\mathbf{b} = [b_1 \ b_2 \ \dots \ b_N]'$ e $\mathbf{k} = [k_1 \ k_2 \ \dots \ k_N]'$, sendo N o número total de animais.

μ_a, μ_b e μ_k são as médias fenotípicas dos parâmetros a , b e k respectivamente; \mathbf{Z} é a matriz de incidência dos efeitos aleatórios com dimensão $3N \times 3q$, sendo q o número de animais com informações no pedigree;

\mathbf{u}_a , \mathbf{u}_b e \mathbf{u}_k são os vetores de valores genéticos de cada parâmetro, com

$$\mathbf{u}_a = [u_{a_1} \ u_{a_2} \ \dots \ u_{a_N}]', \mathbf{u}_b = [u_{b_1} \ u_{b_2} \ \dots \ u_{b_N}]' \text{ e } \mathbf{u}_k = [u_{k_1} \ u_{k_2} \ \dots \ u_{k_N}]';$$

\mathbf{R} é a matriz 3×3 de (co)variância residual, que é definida da seguinte forma:

$$\mathbf{R} = \begin{bmatrix} \sigma_r^2(a) & \sigma_r(a,b) & \sigma_r(a,k) \\ \sigma_r(a,b) & \sigma_r^2(b) & \sigma_r(b,k) \\ \sigma_r(a,k) & \sigma_r(b,k) & \sigma_r^2(k) \end{bmatrix};$$

\mathbf{I} é uma matriz identidade $N \times N$.

Os vetores de valores genéticos \mathbf{u}_a , \mathbf{u}_b e \mathbf{u}_k foram simulados a partir de uma distribuição normal multivariada, tal que:

$$\begin{pmatrix} \mathbf{u}_a \\ \mathbf{u}_b \\ \mathbf{u}_k \end{pmatrix} \sim N \begin{pmatrix} 0 \\ 0, \mathbf{G} \otimes \mathbf{A} \\ 0 \end{pmatrix}$$

em que :

\mathbf{G} é matriz 3×3 , de (co)variância genética que é definida da seguinte forma:

$$\mathbf{G} = \begin{bmatrix} \sigma_g^2(a) & \sigma_g(a,b) & \sigma_g(a,k) \\ \sigma_g(a,b) & \sigma_g^2(b) & \sigma_g(b,k) \\ \sigma_g(a,k) & \sigma_g(b,k) & \sigma_g^2(k) \end{bmatrix};$$

A é a matriz de parentesco entre os animais.

Na segunda etapa os dados de peso de cada animal (y_{ij}) foram simulados considerando-se de uma distribuição normal, sendo que,

$$y_{ij} | a_i, b_i, k_i, \sigma_e^2 \sim N \left\{ a_i \times \left[1 - b_i \exp(-k_i t_{ij}) \right]^3, \sigma_e^2 \right\}$$

em que a_i, b_i e k_i são parâmetros da curva de crescimento gerados na etapa anterior, para o animal i , t_{ij} refere-se a idade em que o peso y_{ij} foi observado e σ_e^2 é a variância do erro, sendo $i = 1, 2, \dots, N$ e $j = 1, 2, \dots, n$, em que n é o número de medidas por animal.

A população necessária à formação da matriz A foi definida considerando três gerações sucessivas não sobrepostas, tal que cada uma delas continha 15 machos e 135 fêmeas, perfazendo o total de 150 animais por geração e, conseqüentemente, 450 animais na população. Foi construída de forma que, na primeira geração, os pais e as mães dos 150 primeiros animais foram considerados desconhecidos. Na segunda geração, os filhos da primeira geração foram os pais da segunda geração e, na terceira geração, os filhos da segunda geração foram os pais da terceira geração. Os acasalamentos entre os animais foram feitos de forma aleatória. Como havia menos fêmeas (135) do que número de filhos (150) nas gerações dois e três, 15 fêmeas foram escolhidas aleatoriamente, nestas gerações, para serem mães de dois animais e todos os machos destas gerações possuíam 10 progênies cada um.

As médias μ_a, μ_b e μ_k foram de 700, 0,4 e 0,003 para os parâmetros a, b e k , respectivamente. As variâncias genéticas e ambientais para cada

parâmetro foram consideradas iguais, assumindo os seguintes valores, $\sigma_g^2(a) = \sigma_r^2(a) = 5000$, $\sigma_g^2(b) = \sigma_r^2(b) = 3,25 \times 10^{-4}$ e $\sigma_g^2(k) = \sigma_r^2(k) = 1,25 \times 10^{-7}$ de acordo com Varona et al.(1999). Assim, a herdabilidade utilizada foi de 0,5 para cada parâmetro. As correlações genéticas e residuais foram assumidas nulas. Foram geradas 15 pesagens por animal com σ_e^2 de 100 kg². O processo de simulação foi repetido 18 vezes.

3.2 Dados observados

Os dados utilizados neste trabalho são oriundos de machos e fêmeas da raça Nelore tradicional, nascidos entre os anos de 2001 a 2004 em dois rebanhos da região de Teófilo Otoni-MG. Estes dados fazem parte dos arquivos de registros da Associação Brasileira de Criadores de Zebu - ABCZ, com sede em Uberaba (MG). Eles se referem a pesagens teoricamente trimestrais (do nascimento aos 500 dias), tendo em vista que elas não ocorriam exatamente a cada três meses. Para que o ajuste das funções de crescimento não fosse comprometido por irregularidades das pesagens dos dados, animais com menos de sete pesagens foram excluídos do estudo. Não houve a separação destes animais por grupos contemporâneos. Assim, animais nascidos em diferentes condições de manejo e época foram todos considerados do mesmo grupo.

3.3 Processo de estimação

Um modelo hierárquico em três estágios foi utilizado para a estimação dos componentes de (co)variância. O primeiro estágio é representado pela modelagem individual da forma da curva de crescimento de cada animal, sendo

este descrito pelo modelo de crescimento de Von Bertalanffy, que é definido por:

$$\dot{y}_{ij} = a_i \left[1 - b_i \exp(k_i t_{ij}) \right]^3 + e_{ij},$$

em que:

y_{ij} é o peso do indivíduo i tomado no tempo j ; ($i = 1, 2, \dots, N$; $j = 1, 2, \dots, n$)

a_i, b_i e k_i são os parâmetros específicos de cada indivíduo i ;

t_{ij} é o tempo correspondente à observação y_{ij} ;

e_{ij} é o erro aleatório, independentemente e normalmente distribuído, com média 0 e variância σ_e^2 .

O segundo estágio da hierarquia descreve como os parâmetros variam entre indivíduos. Cada parâmetro do modelo é determinado por um efeito fixo μ e por um efeito genético u , além de um efeito aleatório não controlado denotado por ε . Considerando essas definições, pode-se utilizar um modelo linear misto para representar a estrutura desses efeitos que, na forma matricial, é dado por:

$$p = X\mu + Zu + \varepsilon$$

em que:

p é um vetor de parâmetros $3N \times 1$, tal que $p = [a_1 \ a_2 \ \dots \ a_N \ b_1 \ b_2 \ \dots \ b_N \ k_1 \ k_2 \ \dots \ k_N]'$;

X é uma matriz, $3N \times 3N$, de incidência dos efeitos fixos;

μ é um vetor, $3N \times 1$, de efeitos fixos;

Z é uma matriz, $3N \times 3q$, de incidência dos efeitos genéticos aditivos,

u é um vetor, $3q \times 1$, de efeitos genéticos aditivos, tal que $u \sim N(0, G \otimes A)$;

ε é um vetor de resíduos referentes ao modelo, tal que $\varepsilon \sim N(0, R \otimes I)$.

O terceiro estágio introduzido na estrutura hierárquica consiste na descrição das distribuições *a priori* a cerca de todos os parâmetros considerados no primeiro e segundo estágios. Para μ , u e p foram atribuídas distribuições normais multivariadas, para as matrizes R e G distribuições de wishart invertida e para σ_e^2 distribuição uniforme, como mostrado a seguir (Chang et al.,2001 Blasco et al., 2003):

$$f(\mu) \propto |V|^{-1/2} \exp\left[-\frac{1}{2}(\mu - m)' V^{-1}(\mu - m)\right]$$

$$f(R) \propto |R|^{-(n_r + d + 1)/2} \exp\left[-\frac{1}{2}tr(n_r R^{-1} S_R^2)\right]$$

$$f(p | \mu, u, R) \propto |R|^{-N/2} \exp\left[-\frac{1}{2}(p - X\mu - Zu)' (R \otimes I)^{-1} (p - X\mu - Zu)\right]$$

$$f(u | G) \propto |G|^{-q/2} \exp\left[-\frac{1}{2}u' (G \otimes A)^{-1} u\right]$$

$$f(G) \propto |G|^{-(n_g + d + 1)/2} \exp\left[-\frac{1}{2}tr(n_g G^{-1} S_G^2)\right]$$

$$\sigma_e^2 \sim U[0, \sigma_{\max}^2]$$

Os termos m , V , S_R^2 , n_r , S_G^2 e n_g são hiperparâmetros, ou seja, valores assumidos para representar a forma de determinadas distribuições *a priori*; d é o número de parâmetros da curva de crescimento ($d = 3$) e σ_{\max}^2 é constante. A matriz A foi obtida utilizando-se o Proc inbreed do programa SAS (1996), a qual requer somente a especificação dos indivíduos de forma colunada, respeitando-se a seguinte ordem: progênie, matriz, reprodutor.

A função de verossimilhança é dada por:

$$f(y_i | p) \propto (\sigma_e^2)^{-\sum_{i=1}^N n_i/2} \exp\left[-\frac{1}{2\sigma_e^2} (y_i - a_i \mathbf{h}_i)' (y_i - a_i \mathbf{h}_i)\right]$$

em que $h_i = [1 - b_i \exp(-k_i t_{ij})]^3$.

Fazendo-se o produto da verossimilhança com as distribuições *a priori* tem-se a distribuição conjunta *a posteriori*, que é representada da seguinte forma:

$$f(p, \mu, u, G, R, \sigma_e^2 | y_i) \propto f(y_i | p) f(p | \mu, u, R) f(\mu) \times \\ \times f(u | G) f(G) f(R) f(\sigma_e^2)$$

em que: y_i denota o vetor de dados amostrais para cada animal i .

Devido à dificuldade de obtenção das marginais, foram obtidas as condicionais *a posteriori* para os parâmetros. Para aquelas que são representadas por distribuições conhecidas, foi utilizado o algoritmo Gibbs Sampler. Para aquelas que não apresentarem distribuições conhecidas, foi utilizado o algoritmo Metropolis-Hastings.

As distribuições condicionais *a posteriori* utilizadas para cada parâmetro estão descritas a seguir. As matrizes de (co)variância ambiental (R) e genética (G) seguem distribuições de Wishart invertida, segundo Sorensen (1996), ou seja,

$$G \sim IW_3 \left\{ (U' A^{-1} U + n_G S_G^2)^{-1}, n_G + q \right\} \text{ e } R \sim IW_3 \left\{ (E' I_N E + n_R S_R^2)^{-1}, n_R + N \right\}$$

em que:

$$U' A^{-1} U = \begin{bmatrix} u_a \\ u_b \\ u_c \end{bmatrix} A^{-1} \begin{bmatrix} u_a & u_b & u_c \end{bmatrix} \text{ e } E' A^{-1} E = \begin{bmatrix} e_a \\ e_b \\ e_c \end{bmatrix} I_N \begin{bmatrix} e_a & e_b & e_c \end{bmatrix}$$

Seguindo as definições de Chang et al. (2003), o parâmetro a e os vetores μ e u são normalmente distribuídos, ou seja,

$$a_i | b_i, k_i, \mu, u, R, G, y \sim N \left[\hat{a}_i, (h_i \sigma_e^{-2} h_i + \Gamma^{aa})^{-1} \right],$$

em que $\hat{a}_i = \frac{\mathbf{h}_i' \sigma_e^{-2} \mathbf{y}_i + r^{aa} m_{a,i} - r^{ab} (b_i - m_{b,i}) - r^{ak} (k_i - m_{k,i})}{\mathbf{h}_i' \sigma_e^{-2} \mathbf{h}_i + r^{aa}}$,

$$\boldsymbol{\mu} | \mathbf{u}, \mathbf{R}, \mathbf{G}, \mathbf{p} \sim N \left[\begin{array}{l} (\mathbf{X}'\mathbf{R}^{-1} \otimes \mathbf{I}\mathbf{X} + \mathbf{V}^{-1})^{-1} \mathbf{X}'\mathbf{R}^{-1} \otimes \mathbf{I}(\mathbf{p} - \mathbf{Z}\mathbf{u}), \\ (\mathbf{X}'\mathbf{R}^{-1} \otimes \mathbf{I}\mathbf{X} + \mathbf{V}^{-1})^{-1} \end{array} \right]$$

$$\mathbf{u} | \boldsymbol{\mu}, \mathbf{R}, \mathbf{G}, \mathbf{p} \sim N \left[\begin{array}{l} (\mathbf{Z}'\mathbf{R}^{-1} \otimes \mathbf{I}\mathbf{Z} + \mathbf{G}^{-1} \otimes \mathbf{A}^{-1})^{-1} \mathbf{Z}'\mathbf{R}^{-1} \otimes \mathbf{I}(\mathbf{p} - \mathbf{X}\boldsymbol{\mu}), \\ (\mathbf{Z}'\mathbf{R}^{-1} \otimes \mathbf{I}\mathbf{Z} + \mathbf{G}^{-1} \otimes \mathbf{A}^{-1})^{-1} \end{array} \right]$$

Para a variância do erro (σ_e^2), foi obtida uma gama inversa, isto é

$$\sigma_e^2 \sim \text{GI} \left[N/2 - 1, \frac{1}{2} \sum_i \sum_j \{y_{ij} - a_i [1 - b_i \exp(-k_i t_{ij})]^3\}^2 \right]$$

Para os parâmetros b e k , as distribuições obtidas não são representadas por distribuições conhecidas, como mostrado abaixo:

$$b_i | a_i, k_i, \boldsymbol{\mu}, \mathbf{u}, \mathbf{R}, \mathbf{G}, \mathbf{y} \sim f(b_i) \text{ e } k_i | a_i, b_i, \boldsymbol{\mu}, \mathbf{u}, \mathbf{R}, \mathbf{G}, \mathbf{y} \sim f(k_i)$$

em que:

$$f(b_i) = \exp\left\{-\frac{1}{2\sigma_e^2} (\mathbf{y}_i - a_i \mathbf{h}_i)' (\mathbf{y}_i - a_i \mathbf{h}_i) - \frac{1}{2r^{bb}} [r^{bb} b_i - (r^{bb} m_{b,i} - r^{ab} (a_i - m_{a,i})) - r^{bk} (k_i - m_{k,i})]^2\right\}$$

$$f(k_i) = \exp\left\{-\frac{1}{2\sigma_e^2} (\mathbf{y}_i - a_i \mathbf{h}_i)' (\mathbf{y}_i - a_i \mathbf{h}_i) - \frac{1}{2r^{kk}} [r^{kk} k_i - (r^{kk} m_{k,i} - r^{ak} (a_i - m_{a,i}) - r^{bk} (b_i - m_{b,i}))]^2\right\}$$

Assim, para os parâmetros a , \mathbf{G} , \mathbf{R} , $\boldsymbol{\mu}$, \mathbf{u} e σ_e^2 , foi utilizado o algoritmo Gibbs Sampler e, para os parâmetros b e k , o algoritmo Metropolis-Hastings. Para a implementação do algoritmo Metropolis-Hastings, utilizou-se a

técnica denominada *Random Walk Chain*, para a obtenção dos valores candidatos. Essa técnica foi utilizada para que os valores de b e k ficassem sempre dentro de seus intervalos paramétricos $[0,1]$.

Segundo Lima (2005), na obtenção de valores candidatos, b^* e k^* , a partir de b_i e k_i , são utilizadas cadeias de passeios aleatórios (*Random Walk Chain*). Dessa forma, um novo valor para o parâmetro é gerado a partir do valor atual, somando-se um valor de uma variável aleatória W , tal que $W \sim N(0, \sigma^2)$, em que σ^2 é um hiperparametro. O processo para a geração de valores candidatos obedece aos seguintes passos:

- i) gera-se um valor w_i , sendo que $W \sim N(0, \sigma_b^2)$ e calcula-se o valor candidato para b_{i+1} , a partir de b_i , que é dado por:

$$\log it(b^*) = \log it(b_i) + w_i \text{ e assim } b^* = \left(1 + \frac{1 - b_i}{b_i \exp w_i}\right);$$

ii) similar para $c^* = \left(1 + \frac{1 - c_i}{c_i \exp w_i}\right)$.

Assim, os valores gerados, b^* e k^* , são os valores candidatos a serem aceitos ou não pelo algoritmo Metropolis-Hastings. A taxa de aceitação desses valores deve permanecer entre 17% e 45%, como indicado por Blasco et al. (2003).

Os algoritmos Gibbs Sampler e Metropolis-Hastings foram implementados no software R. Foram executada 51.000 iterações, respeitando-se, respectivamente, um *burn-in* e um intervalo de amostragem *thin* de 1.000 e 10 iterações. Assim, obteve-se ao final, uma cadeia de 5.000 iterações. Tratando-se de processos iterativos, a constatação de sua convergência é imprescindível para a validade do estudo. Portanto, para verificar a convergência utilizaram-se os

critérios de Geweke (1992) e de Raftery e Lewis (1992), disponíveis no pacote BOA (Bayesian Output Analysis) do aplicativo R.

As distribuições marginais para as herdabilidades (h^2), correlações ambientais (r_r) e correlações genéticas (r_g) foram obtidas da combinação dos componentes de (co)variância estimados, os quais a cadeia convergiu (Rosa, 1998). São representadas da seguinte forma respectivamente:

$$h_a^2 = \frac{\sigma_g^2(a)}{\sigma_g^2(a) + \sigma_r^2(a)}, h_b^2 = \frac{\sigma_g^2(b)}{\sigma_g^2(b) + \sigma_r^2(b)} \text{ e } h_k^2 = \frac{\sigma_g^2(k)}{\sigma_g^2(k) + \sigma_r^2(k)}$$

$$r_r(a,b) = \frac{\sigma_r(a,b)}{\sqrt{\sigma_r^2(a)\sigma_r^2(b)}}, r_r(a,k) = \frac{\sigma_r(a,k)}{\sqrt{\sigma_r^2(a)\sigma_r^2(k)}} \text{ e}$$

$$r_r(b,k) = \frac{\sigma_r(b,k)}{\sqrt{\sigma_r^2(b)\sigma_r^2(k)}}$$

$$r_g(a,b) = \frac{\sigma_g(a,b)}{\sqrt{\sigma_g^2(a)\sigma_g^2(b)}}, r_g(a,k) = \frac{\sigma_g(a,k)}{\sqrt{\sigma_g^2(a)\sigma_g^2(k)}} \text{ e}$$

$$r_g(b,k) = \frac{\sigma_g(b,k)}{\sqrt{\sigma_g^2(b)\sigma_g^2(k)}}$$

4 RESULTADOS E DISCUSSÃO

4.1 Resultados obtidos para os dados simulados

Na Tabela 2 são apresentadas as estimativas (médias *a posteriori*) dos componentes de variância genéticas e ambientais dos parâmetros a , b e k , das herdabilidades de cada parâmetro e das correlações genéticas e residuais entre os

parâmetros. Os valores apresentados são as médias das estimativas pontuais (médias) das 18 repetições do processo de simulação.

TABELA 2 Valores médios obtidos para as estimativas dos componentes de variância e das herdabilidades e correlações genéticas e residuais das repetições do processo de simulação

Variâncias e parâmetros genéticos	Simulado	Média	Desvio padrão
$\sigma_g^2(a)$	5000	5684,93	243,53
$\sigma_g^2(b)$	$3,25 \times 10^{-4}$	$3,42 \times 10^{-4}$	$9,99 \times 10^{-6}$
$\sigma_g^2(k)$	$1,25 \times 10^{-7}$	$1,42 \times 10^{-7}$	$1,48 \times 10^{-8}$
$\sigma_r^2(a)$	5000	4528,52	224,32
$\sigma_r^2(b)$	$3,25 \times 10^{-4}$	$3,14 \times 10^{-4}$	$5,25 \times 10^{-7}$
$\sigma_r^2(k)$	$1,25 \times 10^{-7}$	$1,34 \times 10^{-7}$	$3,89 \times 10^{-10}$
h_a^2	0,5	0,55	0,002
h_b^2	0,5	0,52	0,007
h_k^2	0,5	0,52	0,007
$r_g(a,b)$	0	0,017	0,025
$r_g(a,k)$	0	-0,006	0,02
$r_g(b,k)$	0	0,008	0,01
$r_r(a,b)$	0	0,08	0,09
$r_r(a,k)$	0	-0,15	0,18
$r_r(b,k)$	0	-0,02	0,05

$\sigma_g^2(a), \sigma_g^2(b), \sigma_g^2(k), \sigma_r^2(a), \sigma_r^2(b), \sigma_r^2(k), h_a^2, h_b^2$ e h_k^2 : variâncias genéticas, residuais e herdabilidades para os parâmetros a, b e k ;

$r_g(a,b)$ e $r_r(a,b)$: correlação genética e residual entre os parâmetros a e b ;

$r_g(a,k)$ e $r_r(a,k)$: correlação genética e residual entre os parâmetros a e k ;

$r_g(b,k)$ e $r_r(b,k)$: correlação genética e residual entre os parâmetros b e k .

De acordo com os resultados da Tabela 2, verifica-se que o processo de simulação foi eficiente. Todos os valores médios estimados para os componentes de variância ficaram próximos dos valores simulados. Os valores encontrados para as herdabilidades dos três parâmetros ficaram também bem próximos aos simulados, sendo a maior diferença relativa ao parâmetro a . Em relação às correlações, todos os valores estimados foram bem pequenos, muito próximos de zero, como esperado. Todos esses resultados concordam com os de Varona et al. (1999), que também encontraram valores estimados muito próximos dos simulados, sendo até melhores do que os encontrados no presente estudo. Isso se deve, provavelmente, ao uso de condicionais diferentes no processo de simulação, pois os autores usaram normais univariadas e, neste trabalho, foram utilizadas normais multivariadas.

Em relação à constatação da convergência da simulação em geral, os valores do escore Z foram todos menores que 1,96, como indicado por Blasco et al. (2003), o que, segundo Geweke (1992), significa estacionaridade da cadeia gerada. Além disso, os valores do FD (fator de dependência) de Raftery e Lewis permaneceram praticamente todos menores que 5, o que, segundo Nogueira et al. (2004), significa que a convergência já foi atingida. O critério de Raftery Lewis fornece também as estimativas de *burn-in* e *thin* que devem ser usadas no processo de geração das cadeias. Todos os valores de *burn-in* e *thin* utilizados na simulação foram bem acima daqueles recomendados por esse critério. Outro indicativo de qualidade de convergência são os valores encontrados para os EMC (Erro de Monte Carlo), que foram de baixa magnitude, o que, segundo Blasco et al. (2003), também assegura convergência. Na Tabela 3 está ilustrado como a simulação se comportou em relação à convergência, para a repetição número 8.

TABELA 3 Média, desvio padrão (DP), fator de dependência para o teste de Raftery e Lewis (FD), valor de z para o teste de Geweke (Z) e Erro de Monte Carlo (EMC) obtidas da densidade a *posteriori* das variâncias genéticas e ambientais, herdabilidades, correlações genéticas e ambientais para a repetição 8 do processo de simulação.

Variâncias e parâmetros genéticos	Médias	DP	FD	Z	EMC
$\sigma_g^2(a)$	6028,26	85,39	1,01	1,20	12,09
$\sigma_g^2(b)$	$3,52 \times 10^{-4}$	$7,73 \times 10^{-6}$	5,44	0,36	$4,00 \times 10^{-7}$
$\sigma_g^2(k)$	$1,56 \times 10^{-7}$	$3,43 \times 10^{-9}$	4,96	0,45	$1,05 \times 10^{-9}$
$\sigma_r^2(a)$	4825,86	683,29	1,03	-0,38	12,96
$\sigma_r^2(b)$	$3,15 \times 10^{-4}$	$2,61 \times 10^{-5}$	0,94	-1,73	$4,44 \times 10^{-7}$
$\sigma_r^2(k)$	$1,34 \times 10^{-7}$	$1,21 \times 10^{-8}$	1,10	0,05	$1,05 \times 10^{-9}$
h_a^2	0,55	0,040	0,96	1,24	$6,25 \times 10^{-4}$
h_b^2	0,52	0,022	2,30	1,65	$4,97 \times 10^{-4}$
h_k^2	0,53	0,023	2,41	0,58	$4,86 \times 10^{-4}$
$r_g(a,b)$	$1,44 \times 10^{-2}$	0,085	0,96	1,06	$1,26 \times 10^{-3}$
$r_g(a,k)$	$-1,96 \times 10^{-2}$	0,087	1,03	-0,85	$1,21 \times 10^{-3}$
$r_g(b,k)$	$7,50 \times 10^{-3}$	0,10	1,01	-0,87	$1,86 \times 10^{-3}$
$r_r(a,b)$	$7,78 \times 10^{-2}$	0,091	1,04	-1,48	$1,40 \times 10^{-3}$
$r_r(a,k)$	-0,14	0,089	0,99	0,47	$1,14 \times 10^{-3}$
$r_r(b,k)$	$-2,17 \times 10^{-2}$	0,14	1,97	0,37	$3,06 \times 10^{-3}$

$\sigma_g^2(a), \sigma_g^2(b), \sigma_g^2(k), \sigma_r^2(a), \sigma_r^2(b), \sigma_r^2(k), h_a^2, h_b^2$ e h_k^2 : variâncias genéticas, residuais e herdabilidades para os parâmetros a, b e k ;

$r_g(a,b)$ e $r_r(a,b)$: correlação genética e residual entre os parâmetros a e b ;

$r_g(a,k)$ e $r_r(a,k)$: correlação genética e residual entre os parâmetros a e k ;

$r_g(b,k)$ e $r_r(b,k)$: correlação genética e residual entre os parâmetros b e k .

As Figuras 1, 2 e 3 referem-se aos HPD (intervalo de maior densidade *a posteriori*) das herdabilidades dos parâmetros a e k e da correlação genética entre eles.

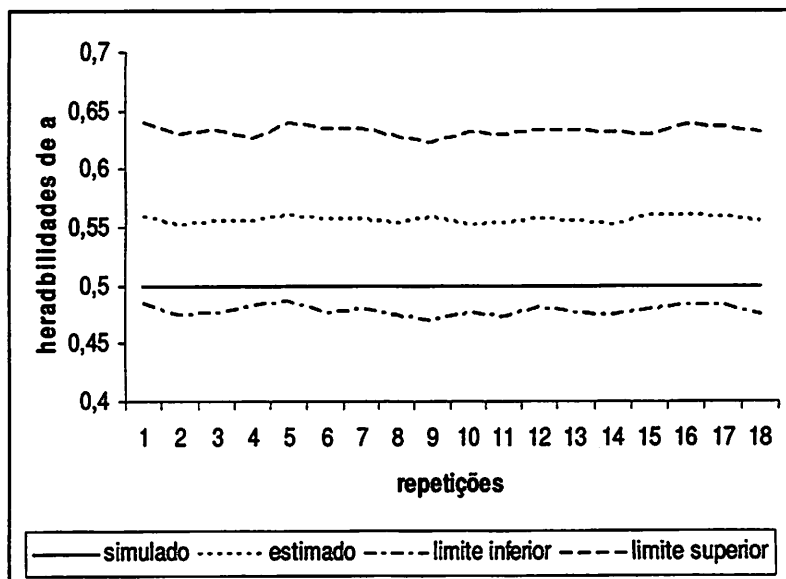


FIGURA 1-Valores paramétricos, estimados e limite inferior e superior do HPD para a herdabilidade do parâmetro a .

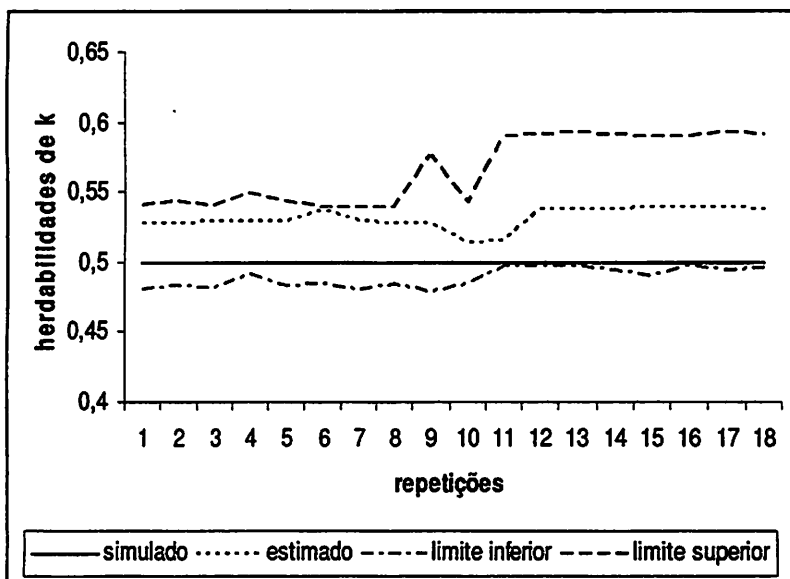


FIGURA 2- Valores paramétricos, estimados e limite inferior e superior do HPD para a herdabilidade do parâmetro k .

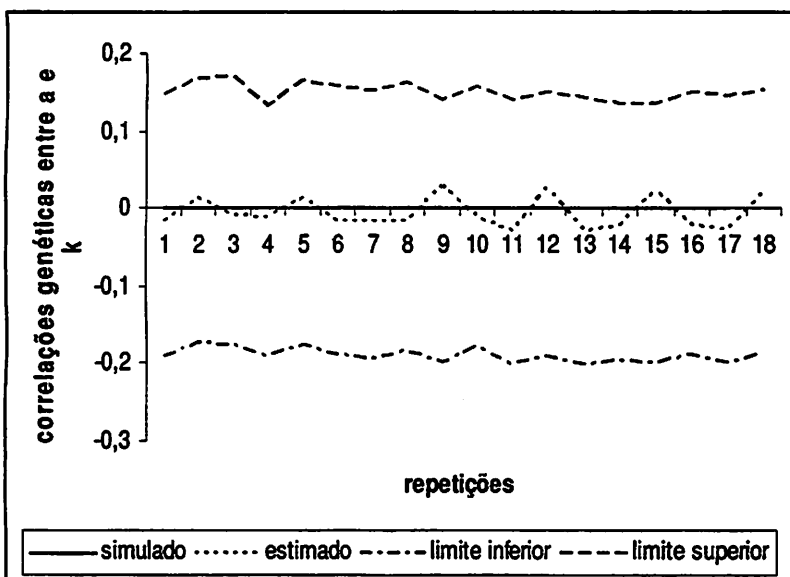


FIGURA 3- Valores paramétricos, estimados e limite inferior e superior do HPD para a correlação genética entre os parâmetros a e k .

Em relação aos resultados mostrados nas Figuras 1, 2 e 3, pode-se afirmar que a metodologia utilizada foi eficiente, pois todos os valores paramétricos considerados na simulação ficaram incluídos no HPD. Para a herdabilidade do parâmetro k , o valor simulado ficou bem próximo ao limite inferior do HPD. Isso se deve, provavelmente, ao fato de k ser um parâmetro que exige mais complexidade para a sua estimação, por ser um parâmetro de condicional não conhecida e necessitar do uso do algoritmo Metropolis-Hastings.

As taxas de aceitação do algoritmo Metropolis-Hastings permaneceram, em média, em 23% para o parâmetro b e em 24% para o parâmetro k , conforme recomendado por Blasco et al. (2003).

Nas Figuras 4 e 5 mostram-se as cadeias, considerando o *burn-in* e *thin*, juntamente com as densidades marginais *a posteriori* para $\sigma_g^2(a)$ e $\sigma_r^2(a)$ para as repetições 1 e 8 do processo de simulação.

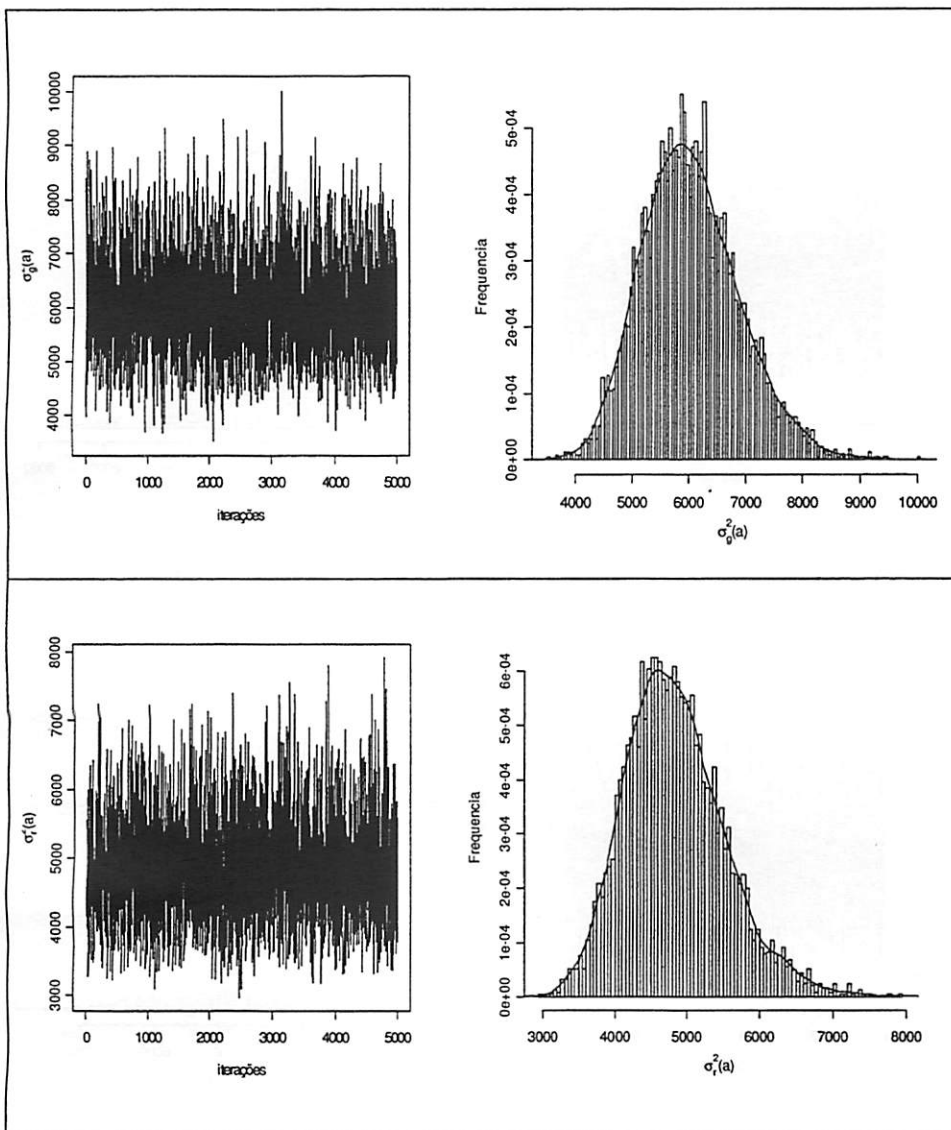


FIGURA 4- Cadeias geradas pelo método MCMC e densidade a posteriori para $\sigma_s^2(a)$ e $\sigma_r^2(a)$ para a repetição 1.

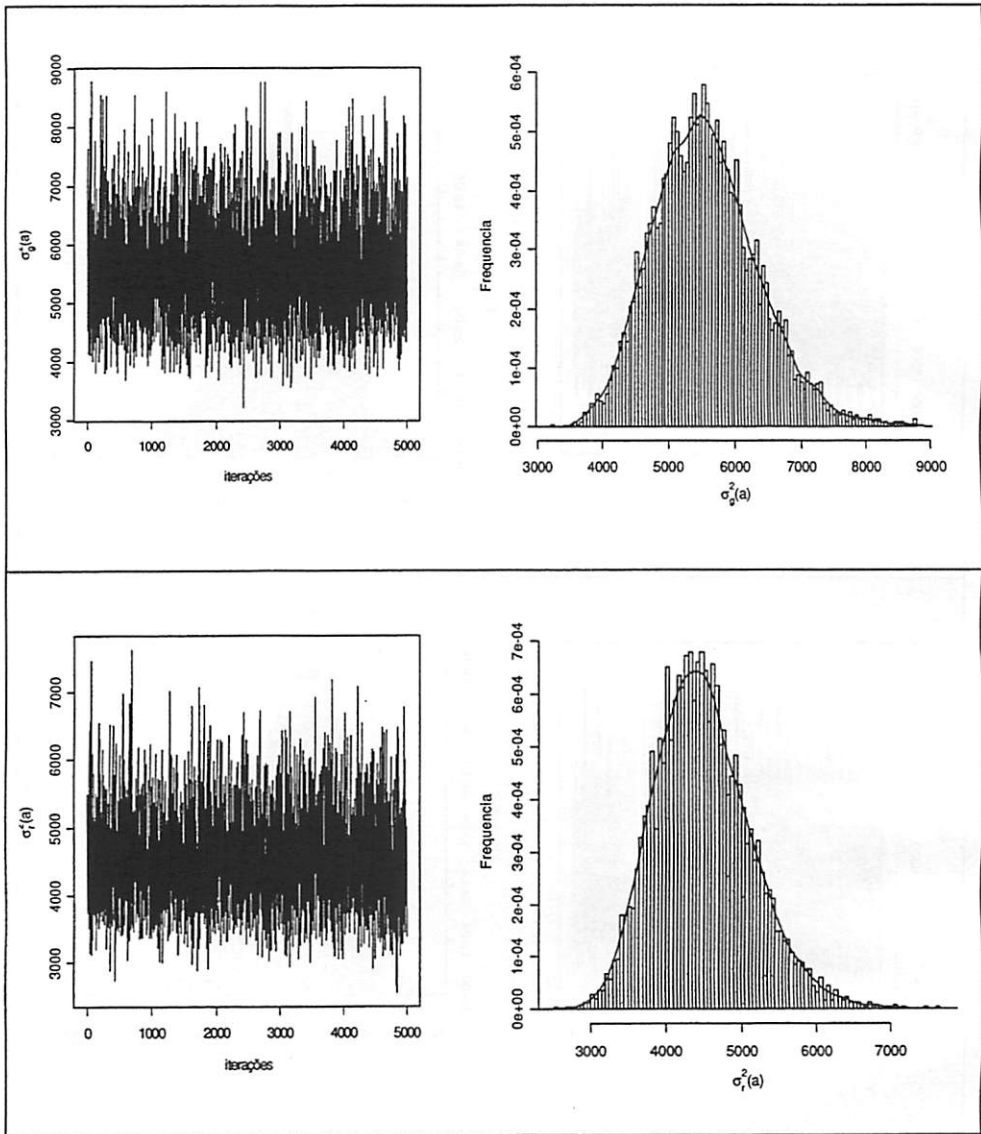


FIGURA 5- Cadeias geradas pelo método MCMC e densidade a posteriori para $\sigma_g^2(a)$ e $\sigma_r^2(a)$ para a repetição 8.

Nas Figuras 4 e 5 mostra-se como $\sigma_g^2(a)$ e $\sigma_r^2(a)$ se comportaram na cadeia final de 5.000 iterações. Conforme o esperado, apresentaram-se sem picos discrepantes e em torno do valor de 5.000. Os resultados mostrados nestas

figuras concordam com os apresentados na Tabela 2, em que a média das repetições se concentrou em torno de 5.600 para $\sigma_g^2(a)$ e 4.500 para $\sigma_r^2(a)$.

4.2 Resultados obtidos para os dados reais

Em relação aos dados reais observa-se na Tabela 4, que a convergência foi alcançada, pois, os valores de Z do teste de Geweke não foram superiores a 1,96, os FD foram menores que 5 e os valores de EMC foram pequenos, o que significa convergência, como foi obtido com os dados simulados. Portanto, nenhuma irregularidade foi observada no processo de amostragem, o que assegura a constatação da convergência requerida.

TABELA 4 Média, desvio padrão (DP), intervalo de maior densidade *a posteriori* (HPD), fator de dependência para o teste de Raftery e Lewis (FD), valor de *z* para o teste de Geweke (Z) e Erro de Monte Carlo (EMC) obtidas da densidade *a posteriori* das variâncias genéticas e ambientais, herdabilidades e correlação genética para os parâmetros *a* e *k*, para os dados reais de Nelore.

Variâncias e parâmetros genéticos	Médias (DP)	HPD		FD	Z	EMC
		LI	LS			
$\sigma_g^2(a)$	2557,60 (484,80)	1647,07	3513,54	1,0	1,1	7,65
$\sigma_r^2(a)$	9748,84 (1044,15)	7700,08	11800,12	0,9	0,1	18,58
$\sigma_g^2(k)$	$1,92 \times 10^{-7}$ ($7,18 \times 10^{-9}$)	$1,78 \times 10^{-7}$	$1,99 \times 10^{-7}$	1,1	0,5	$1,0 \times 10^{-9}$
$\sigma_r^2(k)$	$1,15 \times 10^{-6}$ ($2,66 \times 10^{-7}$)	$8,00 \times 10^{-6}$	$1,66 \times 10^{-6}$	2,1	-0,6	$9,5 \times 10^{-9}$
h_a^2	0,21 (0,03)	0,14	0,26	1,0	1,0	$4,3 \times 10^{-4}$
h_k^2	0,14 (0,02)	0,09	0,19	3,3	0,8	$9,1 \times 10^{-4}$
$r_g(a, k)$	0,02 (0,08)	-0,12	0,19	0,9	-1,3	$1,0 \times 10^{-3}$

$\sigma_g^2(a)$, $\sigma_g^2(k)$, $\sigma_r^2(a)$, $\sigma_r^2(k)$, h_a^2 e h_k^2 : variâncias genéticas, residuais e herdabilidades para os parâmetros *a* e *k*;

$r_g(a, k)$: correlação genética entre os parâmetros *a* e *k*.

Para o parâmetro *a*, a estimativa de herdabilidade encontrada foi de 0,21, que é considerada moderada. Isso indica que este parâmetro poderia ser utilizado para fins de seleção, concordando com Forni (2007), que encontrou alto valor de herdabilidade para o parâmetro *a* (0,52). Elias (1998), em suas revisões, relatou valores de herdabilidade para o parâmetro *a* de animais Nelore, que variaram de 0,13 a 0,55, para o modelo de crescimento de Von Bertalanffy.

Para o parâmetro k , a estimativa da herdabilidade encontrada foi de 0,14 que, apesar de não ser alta, também poderia ser utilizada para fins de seleção. Esse resultado concorda com o de Elias (1998), que relatou valores de herdabilidade variando de 0,12 a 0,44, mas difere do de Forni (2006), que encontrou valores mais baixos para a herdabilidade de k (0,06).

Os valores de herdabilidade encontrados neste trabalho (herdabilidades baixas a moderadas) se devem ao alto valor da variância residual de cada parâmetro neta população. Isso, provavelmente, ocorreu porque os animais foram considerados todos do mesmo grupo contemporâneo. Logo, não houve separação ambiental entre eles no processo de estimação dos parâmetros, o que, provavelmente, fez com que a variância residual ficasse tão alta. Assim, a consideração dos efeitos ambientais no processo de estimação poderia aumentar as estimativas das herdabilidades.

Segundo Elias (1998), observa-se grande variabilidade nas estimativas de herdabilidades dos parâmetros dos vários estudos de crescimento de bovinos. Segundo esse autor, isso é de se esperar, pois os parâmetros obtidos nesses estudos podem ser considerados extrapolações de dados de pesagens obtidas somente até os 24 meses de idade, quando os animais ainda não atingiram a idade adulta. De acordo com Giannoni e Giannoni (1989), a herdabilidade varia de acordo com a espécie animal, com a mesma espécie em diferentes condições, de local para local e entre gerações.

A correlação genética encontrada entre os parâmetros a e k foi de 0,02, que é baixa e positiva, bem próxima de zero, diferindo da maioria dos trabalhos encontrados na literatura, nos quais essa correlação é normalmente alta e negativa. Segundo Santoro et al. (2005), que encontraram correlações genéticas positivas entre os parâmetros a e k para animais Nelore mochos, esse é um comportamento de difícil explicação e, possivelmente, se deve a problemas de amostragem.

Nas Figuras 6 e 7 são apresentadas as cadeias finais e as densidades marginais *a posteriori* para $\sigma_r^2(a)$, $\sigma_g^2(a)$, $\sigma_r^2(k)$ e $\sigma_g^2(k)$ dos dados reais.

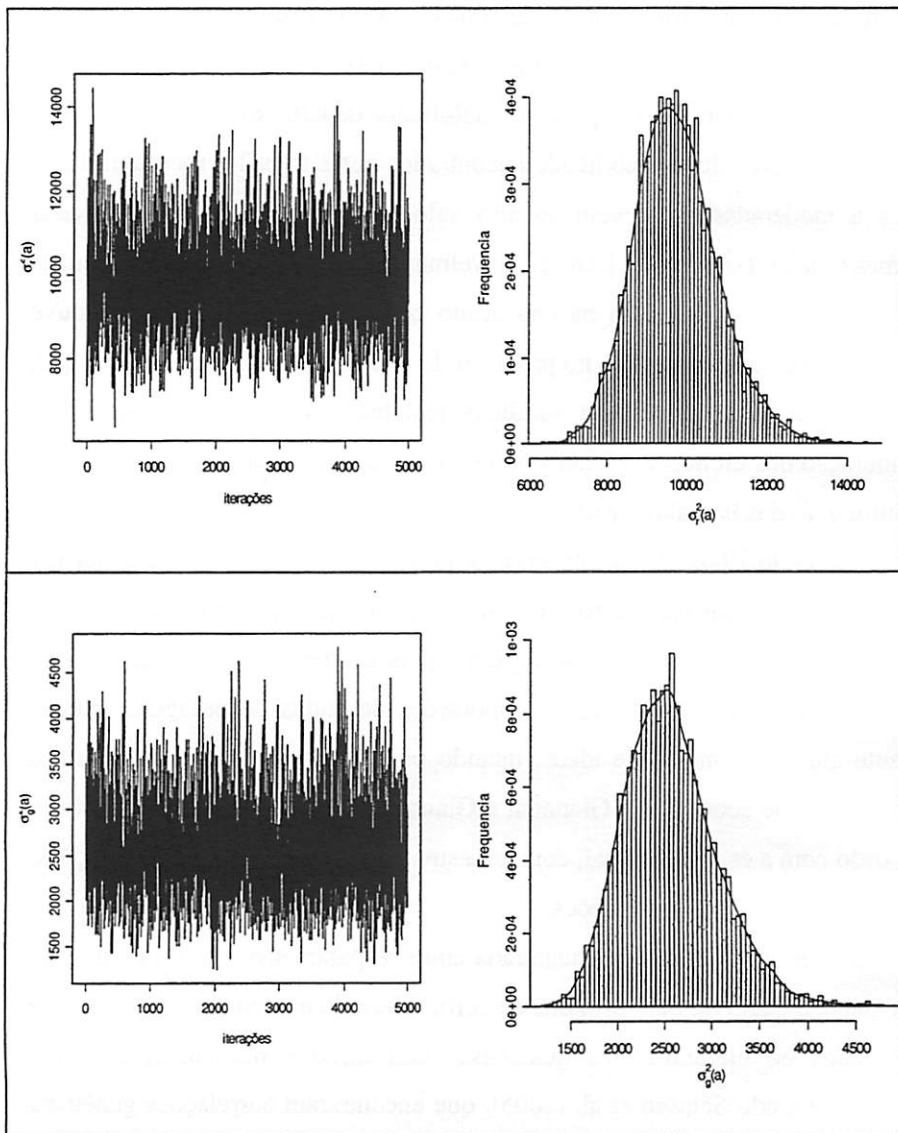


FIGURA 6 Cadeias geradas pelo método MCMC e densidade *a posteriori* para $\sigma_r^2(a)$ e $\sigma_g^2(a)$, para os dados reais de Nelore

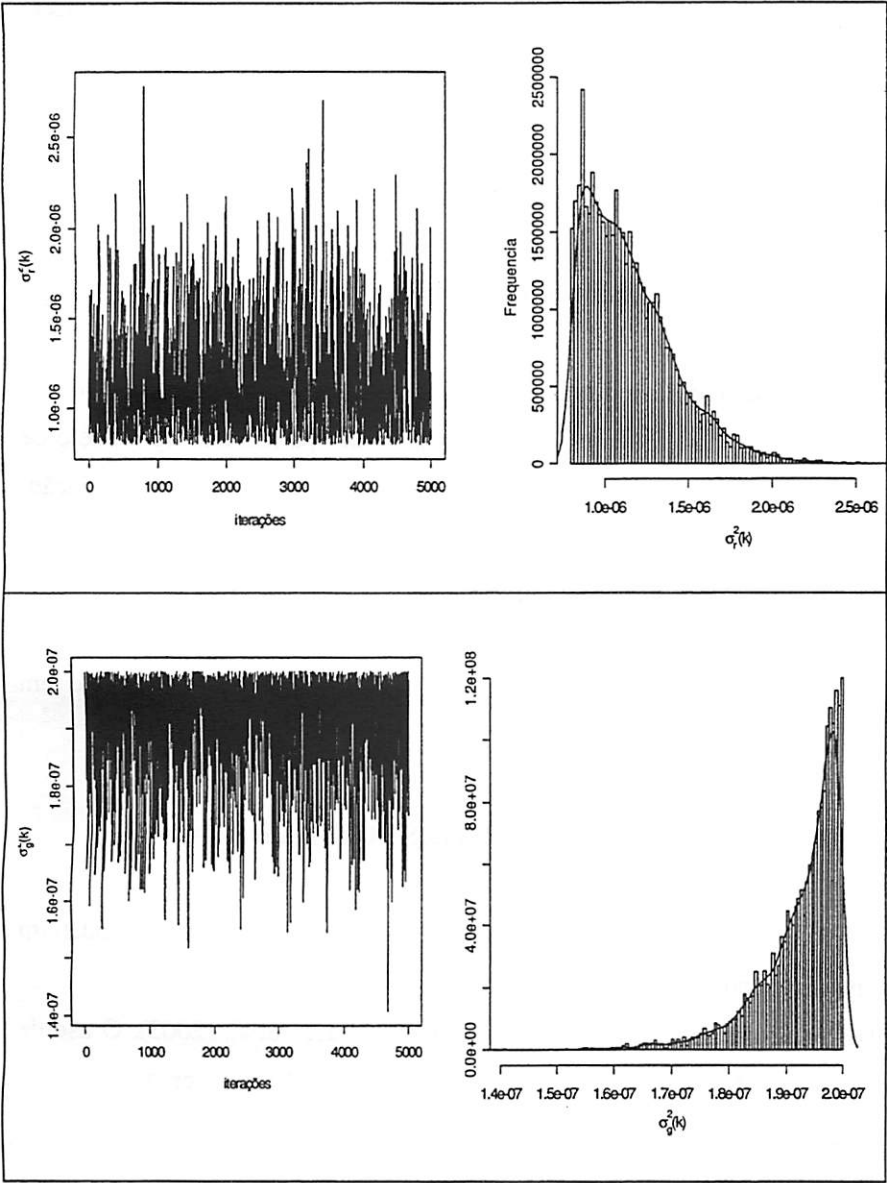


Figura 7 Cadeias geradas pelo método MCMC e densidade *a posteriori* para $\sigma_g^2(k)$ e $\sigma_r^2(k)$ para os dados reais de Nelore

Os resultados mostrados nas Figuras 6 e 7 mostram o comportamento da cadeia final do processo iterativo, confirmando também os resultados

apresentados na Tabela 4, a respeito das médias *a posteriori* de $\sigma_g^2(a)$, $\sigma_r^2(a)$, $\sigma_g^2(k)$ e $\sigma_r^2(k)$, para os dados reais.

5 CONCLUSÕES

Com base neste trabalho pode-se concluir que:

- a metodologia hierárquica bayesiana se mostrou uma alternativa viável para se estimar os componentes de (co)variância dos parâmetros do modelo de Von Bertalanffy, pois produziu boas estimativas do processo de simulação, próximas aos valores simulados e com valores paramétricos sempre dentro dos HPD. Além disso, produziu estimativas para os dados reais condizentes com a realidade;

- os parâmetros a e k podem ser usados para fins de seleção em programas de melhoramento animal.

6- PROPOSTAS FUTURAS

Em estudos futuros, os procedimentos usados neste trabalho poderiam ser complementados. Como por exemplo, poderia ser usado o processo de modelagem da variância do erro assim como fez Blasco et al. (2003). O uso de *prioris* univariadas para os parâmetros também poderiam ser testadas. A utilização de modelos mistos mais complexo também seria uma sugestão para a complementação do trabalho.

7 REFERÊNCIAS BIBLIOGRÁFICAS

BLASCO, A.; PILES, M.; VARONA, L.A. Bayesian analysis of the effect of selection for growth rate on growth curves in rabbits. **Genetics Selection Evolution**, Valencia, v.35, p.21-41, maio 2003.

BROEMELING, L.D. **Bayesian analysis of linear models**. New York: J. Wiley, 1989. 412p.

CARRIJO, S.M.; DUARTE, F.A.M. Description and comparison of growth parameters in Chianina and Nelore cattle breeds. **Genetics and Molecular Biology**, São Paulo, v.22, n.2, p.187-196, jun. 1999.

CHANG, Y.M.; REKAYA, R.; GIANOLA, D.; THOMAS, D.L. Genetic variation of lactation curves in dairy sheep: a Bayesian analysis of Wood's function. **Livestock Production Science**, Amsterdam, v.71, n.2-3, p.241-251, out. 2001.

DeNISE, R.S.K.; BRINKS, J.S. Genetic and environmental aspects of the growth curve parameters in beef cows. **Journal of Animal Science**, v.61, n.6, p.1431-1440, dez. 1985.

ELIAS, M.A. **Análise de curvas de crescimento de vacas das raças Nelore, Guzerá e Gir**. 1998. 128p. Dissertação (Mestrado em Zootecnia) - Escola Superior de Agricultura de "Luiz de Queiroz", Piracicaba, SP.

FARIA, C.U.; MAGNABOSCO, C.U.; ALBUQUERQUE, L.G.; REYS, A.L.; LOBO, L.G.; BEZERRA, L.A.F. Bayesian and frequentist approaches in three trait genetic analysis for growth and reproduction in nelore cattle. In: **WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION**, 8., 2006, Belo Horizonte. **Anais...** Belo Horizonte, 2006. p.13-18.

FARIA, C.U.; MAGNABOSCO, C.U.; ALBUQUERQUE, L.G.; REYS, A.L.; LOBO, L.G.; BEZERRA, L.A.F. Inferência bayesiana e sua aplicação na avaliação genética de bovinos da raça nelore: revisão bibliográfica. **Ciência Animal Brasileira**, Goiania, v.8, n.1, p.75-86, jan./mar. 2007

FITZHUGH, H.A. Analysis of growth curves and strategies for altering their shapes. **Journal of Animal Science**, v.42, n.4, p.1036-1051, jul. 1976.

FORNI, S. Análise da curva de crescimento de bovinos da raça nelore utilizando funções não lineares em análises bayesianas. 2007. 75p. Tese (Doutorado em Zootecnia) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Jaboticabal, SP.

GAMERMAN, D.; LOPES, H.F. Markov Chain Monte Carlo: stochastic simulation for bayesian inference. 2.ed. Londres: Chapman & Hall/CRC, 2006. v. 1, 336p.

GARNERO, A.V.; MARCONDES, C.R.; BEZERRA, L.A.F.; OLIVEIRA, H.N.; LÔBO, H.N. Parâmetros genéticos da taxa de maturação e do peso assintótico de fêmeas da raça Nelore. Arquivo Brasileiro de Medicina Veterinária e Zootecnia, Belo Horizonte, v.57, n.5, p.652- 662, out. 2005.

GELFAND, A.E.; SMITH, A.E.M. Sampling based approaches to calculating marginal densities. Journal of the American Statistical Association, Alexandria, v.85, n.4, p.398-409, jun. 1990.

GIANNONI, M.A.; GIANNONI, M.L. Genética e melhoramento de rebanho nos trópicos. 2.ed. São Paulo: Nobel, 1989. 463p.

GIANNOTTI, J.G.; PACKER, I.U.; MERCADANTE, M.E.Z. Meta-Análise para Estimativas de Herdabilidade para Características de Crescimento em Bovinos de Corte Revista Brasileira de Zootecnia, v.34, n.4, p.1173-1180, 2005.

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, Belo Horizonte, v.57, n.3, p.415-417, jun. 2005.

GEWEKE, J. Evaluating the accuracy of Sampling-based approaches to the calculation of posterior moments. In: BERNARDO, J.M.; BERGER, J.O.; DAWID, A.P.; SMITH, A.F.M. (Ed.). Bayesian statistics. New York: Oxford: Oxford University, 1992. p.625-631.

JEFFREYS, H. Theory of probability. Oxford, UK: Clarendon, 1961. 325p.

LAIRD, A.K.; HOWARD, A. Growth curves in inbred mice. Nature, London, v.213, n.5078, p.786-788, fev. 1967.

LIMA, R.R. Modelagem espaço-temporal para dados de incidência de doenças em plantas. 2005. 149p. Tese (Doutorado em Estatística e Experimentação Agrônômica) - Universidade de São Paulo, Piracicaba, SP.

MENDES, P. N. Curvas de crescimento difásicas de fêmeas hereford com erros auto regressivos e heterogeneidade de variâncias. 2007. 98p. Dissertação (Mestrado em Agronomia) - Universidade Federal de Lavras, Lavras, MG.

MUNIZ, J.A.; SILVA, F.F., GUEDES, M.H.P.; GONÇALVES, T.M. Evaluation of lactation curve of low-yielding gir cows: a Bayesian approach *Acta Scientiarum Animal Science*, Maringá, v.29, n.1, p.79-83, 2007.

NOGUEIRA, D.A.; SÁFADI, T.; FERREIRA, D.F. Avaliação de critérios de convergência univariados para o método de Monte Carlo via Cadeias de Markov. *Revista Brasileira de Estatística*, Rio de Janeiro, v.65, n.224, p.59-88, 2004.

O'HAGAN, A. Kendall's advanced theory of statistics: bayesian inference. New York: E. Arnold, 1994. v.2b, 402p.

OLIVEIRA, H.N.; LÔBO, R.B.; PEREIRA, C.S. Comparação de modelos não-lineares para descrever o crescimento de fêmeas da raça Guzerá. *Pesquisa Agropecuária Brasileira*, Brasília, v.35, n.9, p.1843-1851, set. 2000.

R DEVELOPMENT CORE TEAM. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2007. Disponível em: <<http://www.R-project.org>>. Acesso em: 20 nov. 2007.

RAFTERY A.E.; LEWIS, S. How many iterations in the Gibbs Sampler? In: BERNARDO J.M.; BERGERE, J.O.; DAWID, A P.; SMITH, A.F. (Ed.). *Bayesian statistics*. Oxford: Oxford University, 1992. p.763-773.

ROSA, G.J.M. Análise Bayesiana de modelos Mistos Robustos via Amostrador de Gibbs. 1998. 57p. Tese (Doutorado em Estatística e Experimentação Agrônômica) - Universidade de São Paulo, Piracicaba, SP.

SANTORO, R.S.; BARBOSA, S.B.P.; BRASIL, L.H.A.; SANTOS, E.S. Estimativas de parâmetros de curvas de crescimento de bovinos zebu, criados no estado de Pernambuco. *Revista Brasileira de Zootecnia*, Viçosa, MG, v.34, n.6, p.2262-2279, nov./dez. 2005.

SANTOS, R. **Os cruzamentos na pecuária tropical**. Agropecuária Tropical, 1999. 672p.

SEARLE, S.R.; CASELLA, G.; McCULLOCH, C.E. **Variance componets: J**. Wiley, 1992. 501p.

SILVA, N.A.M.; MUNIZ, J.A.; SILVA, F.F.; AQUINO, L.H.; GONCALVES, T. M. Aplicação do método Bayesiano na estimação de curvas de crescimento de animais da raça Nelore. *Revista Ceres*, Viçosa, v.54, p.192-199, abr./maio 2008a.

SILVA, F.F.; SÁFADI, T.; MUNIZ, J.A.; AQUINO, L.H.; MOURÃO, G.B. Comparação Bayesiana de modelos de previsão para diferenças esperadas nas progênes no melhoramento genético da gado Nelore. *Pesquisa Agropecuária Brasileira*, Brasília, v.43, n.1, p.37-45, jan. 2008b.

SILVA, N.A.; SILVA, F.F.; MUNIZ, J.A.; AQUINO, L.H.T. Estudo de parâmetros de crescimento de bezerros Nelore por meio de um modelo de regressão linear: uma abordagem Bayesiana. *Ciência Animal Brasileira*, v. 7, n. 1, p. 57-65, jan./mar. 2006.

SILVA, F.F.; MUNIZ, J.A.; AQUINO, L.H.; SÁFADI, T. Abordagem Bayesiana da curva de lactação de cabras Saanen de primeira e segunda ordens de parto. *Pesquisa Agropecuária Brasileira*, Brasília, v.40, n.1, p.27-33, jan. 2005.

SILVA, F.F.; AQUINO, L.H.; OLIVEIRA, A.I.G. Estimativas de parâmetros genéticos de curva de crescimento de gado Nelore (*Bos indicus*). *Ciência e Agrotecnologia*, Lavras, v.26, p.1562-1567, 2002. Edição Especial.

SILVA, F.F.; AQUINO, L.H. de; OLIVEIRA, J.A. Influência de fatores genéticos e ambientais sobre as estimativas dos parâmetros das funções de crescimento em gado Nelore. *Ciência e Agrotecnologia*, Lavras, v.25, n.5, p.1195-1205, 2001.

SORENSEN, D. **Gibbs sampling in quantitative genetic** Copenhagen: Foulun, 1996. 186p.

SOUZA, J.C. de; BIANCHINI SOBRINHO, E. Estimativa do peso de bovino de corte, aos 24 meses, da raça Nelore, usando curvas de crescimento. *Revista da Sociedade Brasileira de Zootecnia*, Viçosa, v.23, n.1, p.85-91, jan./fev. 1994.

STATISTICAL ANALYSIS SYSTEM INSTITUTE. SAS/STAT[®]: user's guide. Statistics version 6.4. Cary, 1996. 168p.

STIGLER, S.M. **The history of statistics. The measurement of uncertainty before 1900.** Boston: Harvard University, 1986. 432p.

TALHARI, F.M.; ALENCAR, M.M.; MASCIOLI, A.S.; SILVA, M.A.; BARBOSA, P.F. Correlações genéticas entre características produtivas de fêmeas em um rebanho da raça Canchim **Revista Brasileira de Zootecnia**, Viçosa, v.32, n.4, p.880-886, jul./ago. 2003.

TEDESCHI, L.O.; BOIN, C.; NARDON, R.F.; LEME, P.F. Estudo da curva de crescimento de animais da raça guzerá e seus cruzamentos alimentados a pasto, com e sem suplementação. 1. análise e seleção das funções não-lineares **Revista Brasileira de Zootecnia**, Viçosa, v.29, n.2, p.630-637, mar./abr. 2000.

TORAL, F.L.B.; ALENCAR, M.M.; FREITAS, A.R. Abordagens freqüentista e bayesiana para avaliação genética de bovinos da raça Canchim para características de crescimento **Revista Brasileira de Zootecnia**, Viçosa, v.36, n.1, p.43-53, jan./fev. 2007.

VAN TASSEL, C.P.; CASELLA, G.; POLLAK, E.J. Effects of selection on estimates of variance components using Gibbs sampling and restricted maximum likelihood. **Journal of Dairy Science**, v.78, p.678-692, Sept. 1995.

VARONA, L.; MORENO, C.; GARCIA CORTÉS, L.A.; ALTARRIBA, J. Multiple trait genetic analysis of underlying biological variables of production functions. **Livestock Production Science**, Amsterdam, v.47, p.201-209, Oct. 1997.

VARONA, L.; MORENO, C.; GARCIA CORTÉS, L.A.; YAGÜE, G.; ALTARRIBA, J. Two-step versus joint analysis of Von Bertalanffy function. **Journal of Animal Breeding and Genetics**, Berlin, v.116, p.331-338, jun. 1999.