

UNIVERSIDADE ESTADUAL DE MARINGÁ
CENTRO DE CIÊNCIAS AGRÁRIAS

**ESTIMAÇÃO DE COMPONENTES DE (CO)VARIÂNCIA
E DE TENDÊNCIAS GENÉTICAS EM POPULAÇÕES
SIMULADAS**

Autora: Daniela Andressa Lino
Orientador: Prof. Dr. Elias Nunes Martins

Dissertação apresentada, como parte das exigências para obtenção do título de MESTRE EM ZOOTECNIA, no Programa de Pós-Graduação em Zootecnia da Universidade Estadual de Maringá – Área de Concentração: Produção Animal.

MARINGÁ
Estado do Paraná
Setembro – 2006

UNIVERSIDADE ESTADUAL DE MARINGÁ
CENTRO DE CIÊNCIAS AGRÁRIAS

**ESTIMAÇÃO DE COMPONENTES DE (CO)VARIÂNCIA
E DE TENDÊNCIAS GENÉTICAS EM POPULAÇÕES
SIMULADAS**

Autora: Daniela Andressa Lino
Orientador: Prof. Dr. Elias Nunes Martins

Dissertação apresentada, como parte das exigências para obtenção do título de MESTRE EM ZOOTECNIA, no Programa de Pós-Graduação em Zootecnia da Universidade Estadual de Maringá – Área de Concentração: Produção Animal.

MARINGÁ
Estado do Paraná
Setembro – 2006



UNIVERSIDADE ESTADUAL DE MARINGÁ
CENTRO DE CIÊNCIAS AGRÁRIAS

**ESTIMAÇÃO DE COMPONENTES DE (CO)VARIÂNCIA
E DE TENDÊNCIAS GENÉTICAS EM POPULAÇÕES
SIMULADAS**

Autor: Daniela Andressa Lino

Orientador: Prof. Dr. Elias Nunes Martins

Co-Orientadora: Prof^ª. Dr^ª. Eliane Gasparino

TITULAÇÃO: Doutor em Zootecnia – Área de Concentração Produção
Animal

APROVADA em 02 de março de 2006.

“Mas o caminho dos justos é como a luz do amanhecer, que vai brilhando mais e mais até ser dia perfeito”.

Provérbios, 4:18

A Deus por permitir que tudo isso se tornasse real, pois o homem faz os seus projetos,
mas a resposta vem unicamente de Deus.

Aos meus pais David Lino e Vera Lúcia Lino, por todo o apoio e por acreditarem que
investir no estudo é a maior herança que os pais podem deixar aos filhos.

Ao meu namorado, companheiro, amigo, incentivador e ajudante Jeferson Menezes
Lourenço...

...DEDICO

AGRADECIMENTOS

Agradeço de forma muito especial...

A Deus por permitir que este trabalho fosse realizado.

À Universidade Estadual de Maringá (UEM) e ao Programa de Pós-graduação em Zootecnia (PPZ), por dar oportunidade para a realização deste trabalho.

À CAPES (Coordenação de Aperfeiçoamento do Pessoal de Nível Superior) pelo financiamento e concessão da bolsa de estudos.

Ao Professor Dr. **Elias Nunes Martins**, pela orientação, pelo conhecimento compartilhado, pela amizade, pelo incentivo, por lançar a cada dia um novo desafio e por acreditar que uma Bióloga pode ser Mestre em Zootecnia.

Ao Professor Dr. **Eduardo Shiguero Sakaguti**, pela orientação inicial, pela amizade e por ter me apresentado à linguagem de Programação Fortran.

À Professora Dr^a. **Eliane Gasparino**, pela co-orientação, pela amizade, pelas conversas e risadas e pelo apoio.

Ao amigo e colega de pós-graduação **Carlos Antonio Lopes de Oliveira**, por ceder gentilmente 10 dos seus 360 rebanhos para a realização deste trabalho, pela amizade, pelos ensinamentos na área de melhoramento, pelas lições de vidas, por ser o “debugger” de alguns programas e pela enorme paciência.

Ao Professor Dr. **Luis Otávio Campos Silva**, à amiga **Andréa Gondo** e ao Dr. **Roberto Torres Junior**, ambos da EMPRAPA - gado de corte, pela ajuda nos problemas de limitação computacional, pela amizade e pelo apoio.

Ao amigo Msc. **Renato Elias**, da UFRJ, ao professor **Luis Guilherme Pereira Lima** e ao Professor Dr. **Marcelo Lenzi** pela ajuda e ensinamentos na área computacional.

À amiga e colega de pós-graduação **Alexandra Inês dos Santos**, pela amizade, pelos ensinamentos e por me mostrar que a vida tem diferentes valores.

À amiga **Meiby Carneiro de Paula** e colegas de pós-graduação, pela amizade e pelos ensinamentos. Ao amigo Prof. Dr. **Luis Daniel Giusti Bruno**, pela amizade e apoio.

Aos amigos e colegas de pós-graduação: **Fabiana Martins Costa**, **Ana Paula Ton**, **Ana Carolina Müller Conti**, **Emília de Paiva**, **Priscilla Georg** e **Alexandre Leseur dos Santos**, por trazerem harmonia ao grupo de Melhoramento Genético Animal e pela amizade.

À CTO Construtora e seus proprietários **Edinei** e **Neide Lourenço**, por emprestarem os computadores, e permitir que as análises fossem mais rápidas.

Às amigas de república **Marcela Battilani**, **Denise Maistro** e **Géssica Lourenço**, pela amizade, companheirismo, lições de vida e alegrias compartilhadas.

Às grandes amigas **Fernanda** e **Ana Theodora Della Rosa Rehbein**, por fazerem parte da minha vida.

Aos meus pais **David Lino** e **Vera Lúcia Lino** e à minha irmã **Vanessa Christina Lino**, pelo enorme apoio.

Ao meu namorado **Jeferson Menezes Lourenço**, pelo amor, pela amizade, apoio, ajuda e por estar sempre ao meu lado.

BIOGRAFIA DO AUTOR

Daniela Andressa Lino, filha de David Lino e Vera Lúcia Lino, nasceu na cidade de Astorga, Paraná, em 06 de março de 1982.

Em março de 2000, ingressou no Curso de Ciências Biológicas na Universidade Estadual de Maringá, em Maringá, Paraná e concluiu o curso no ano letivo de 2003, colou grau em maio de 2004 quando obteve o título de Bacharel e Licenciada em Biologia.

Em março de 2004, matriculou-se no curso de Pós-graduação em Zootecnia, em nível de Mestrado, na área de concentração em Produção Animal, na Universidade Estadual de Maringá, realizando pesquisas na área de Melhoramento Genético Animal. Foi, inicialmente, orientada pelo Professor Dr. Eduardo Shiguero Sakaguti, que se ausentou e entregou a orientação ao Professor Dr. Elias Nunes Martins.

Em setembro de 2006, submeteu-se aos exames finais de defesa de dissertação, para se habilitar ao título de Mestre em Zootecnia.

ÍNDICE

	Página
LISTA DE TABELAS.....	ix
LISTA DE FIGURAS.....	x
TABELAS DO APÊNDICE.....	xi
RESUMO.....	xii
ABSTRACT.....	xiv
I – INTRODUÇÃO GERAL E REFERÊNCIAS BIBLIOGRÁFICAS	1
II – OBJETIVOS GERAIS	8
Citação Bibliográfica	9
III – ESTIMAÇÃO DE (CO)VARIÂNCIAS GENÉTICAS HETEROGÊNEAS EM REBANHOS FECHADOS SOB SELEÇÃO	11
Resumo	11
Abstract	12
Introdução	13
Material e Métodos	14
Resultados e Discussão	22
Conclusão	32
Literatura Citada	32

IV – MÉTODOS PARA ESTIMAÇÃO DE TENDÊNCIA GENÉTICA	35
Resumo	35
Abstract	36
Introdução	37
Material e Métodos	38
Resultados e Discussão	49
Conclusão	61
Literatura Citada	62
V – CONCLUSÕES GERAIS	64
VI – APÊNDICE	65

LISTA DE TABELAS

	Página
Tabela 1 – Componentes de variância genética aditiva estimados por geração para P550	24
Tabela 2 – Intervalos de Credibilidade (IC) para as variâncias genéticas de geração de seleção.....	26
Tabela 3 – Valores dos coeficientes da regressão estimados para cada rebanho e coeficientes de determinação (R^2) das equações ajustadas	27
Tabela 1 – Média dos valores genéticos preditos por ano, para a condição 1 (C1).....	50
Tabela 2 – Média dos valores genéticos preditos por ano, para a condição 2 (C2).....	50
Tabela 3 – Mudança genética média em kg por ano, para valor genético real, C1 e C2.....	51
Tabela 4 – Componentes de variância genética aditiva estimados por ano para P550	53
Tabela 5 – Componentes de covariância genética aditiva estimados por ano para P550	53
Tabela 6 – Coeficientes de regressão linear e quadráticos para os métodos TG1, TG2, TG3 e TG4 para a condição 1 (C1)	54
Tabela 7 – Coeficientes de regressão linear e quadráticos para os métodos TG1, TG2, TG3 e TG4 para a condição 2 (C2)	55

LISTA DE FIGURAS

	Página
Figura 1 – Superfície de resposta para (co)variâncias genéticas por ano de nascimento, para P550, para o rebanho 1. O gradiente vermelho até azul é traduzido por valores maiores até valores menores. 1a mostra a redução das variâncias e covariâncias, 1b mostra a redução das variâncias e em 1c a redução das covariâncias	29
Figura 2 – Comportamento das variâncias genéticas por geração de seleção e por ano de nascimento, para P550 dias, para os rebanhos 1 ao 5.....	30
Figura 3 – Comportamento das variâncias genéticas por geração de seleção e por ano de nascimento, para P550 dias, para os rebanhos 6 ao 10.....	31
Figura 1 – Curvas de regressão das médias dos valores genéticos reais, preditos pela avaliação genética e pelos métodos TG1, TG2, TG3 e TG4, em função do ano de nascimento do animal, para os rebanhos de 1 a 5, na condição 1 (C1). Vg é o valor genético	56
Figura 2 – Curvas de regressão das médias dos valores genéticos reais, preditos pela avaliação genética e pelos métodos TG1, TG2, TG3 e TG4, em função do ano de nascimento do animal, para os rebanhos de 6 a 10, na condição 1 (C1). Vg é o valor genético	57
Figura 3 – Curvas de regressão das médias dos valores genéticos reais, preditos pela avaliação genética e pelos métodos TG1, TG2, TG3 e TG4, em função do ano de nascimento do animal, para os rebanhos de 1 a 6, na condição 2 (C2). Vg é o valor genético	58
Figura 4 – Curvas de regressão das médias dos valores genéticos reais, preditos pela avaliação genética e pelos métodos TG1, TG2, TG3 e TG4, em função do ano de nascimento do animal, para os rebanhos de 7 a 10, na condição 2 (C2). Vg é o valor genético	59

TABELAS DO APÊNDICE

	Página
Tabela 1A – Componentes de variância genética aditiva estimados por ano para P550	66
Tabela 2A – Componentes de covariância genética aditiva estimados por ano para P550, para os rebanho de 1 a 5	67
Tabela 3A – Componentes de covariância genética aditiva estimados por ano para P550, para os rebanho de 6 a 10	68

RESUMO

Foram estudados métodos alternativos para a estimação da tendência genética para peso aos 550 dias em 10 rebanhos de bovinos de corte simulados, submetidos à 20 anos de seleção, considerando a existência de parentesco entre os animais e contemplando a heterogeneidade de (co)variância pela inclusão da estrutura de (co)variância genética entre os anos de nascimento. Para a construção da estrutura de (co)variância, foi utilizado um modelo de regressão múltipla, a partir dos componentes estimados para cada geração de seleção, obtidos por análise multicarácter implementada para inferência bayesiana. O modelo de regressão múltipla apresentou, em média, coeficiente de determinação igual a 0,9027 com desvio-padrão de 0,0394, mostrando-se útil na estimação de componentes de (co)variância ao longo dos anos de nascimento dos animais, usados os componentes estimados em cada geração. As tendências genéticas foram estimadas para 8 anos de seleção, com duas condições (C1 e C2) aplicadas à avaliação genética; na C1 foi realizada a avaliação, considerando todos os animais até o 20º ano de seleção, e na C2, apenas os animais que participaram do conjunto de dados utilizado para fazer a estimação da tendência genética, obtida por meio de equações de regressão ajustadas, alternativamente, pelos métodos de quadrados mínimos ordinários (OLS), quadrados mínimos ponderados (WLS) e quadrados mínimos generalizados (GLS). Este último sob duas formas: uma considerando homogeneidade de variância, e a outra, heterogeneidade de variância ao longo dos anos de seleção. Foi verificado comportamento quadrático dos valores genéticos em função do ano de nascimento dos animais, para os 10 rebanhos estudados na condição 1 e para os rebanhos 1, 4, 8 e 9 na condição 2. Para a C1, os métodos WLS e GLS apesar de permitirem maior inclusão de informações no modelo, não foram capazes de detectar as mudanças genéticas com maior precisão, sendo o OLS o método que permitiu maior aproximação às médias

anuais dos valores genéticos reais e preditos. Para a C2, o método GLS, considerando heterogeneidade de variância ao longo dos anos de nascimento do animal, foi o que mais se aproximou das mudanças genéticas reais.

Palavras-chave: componentes de variância, regressão múltipla, heterogeneidade de variância, inferência Bayesiana

ABSTRACT

Alternative methods for the estimation of genetic trends for weight at 550 days (P550) in 10 herds, submitted to 20 years of selection were studied. The familiar relationship between animals and the heterogeneity of (co)variance by the inclusion of genetic (co)variance between the year of birth of the animals were contemplated. For the construction of the (co)variance structure a model of multiple regression from the estimated components for each generation of selection, obtained through multi-character analysis for Bayesian inference, was used. The determination coefficient (R^2) was high with mean of 0.90272 and standard deviation of 0.03936. This indicates that genetic (co)variance components among years can be estimated precisely by multiple regression using the components estimated in each generation. The genetic trends were estimated for 8 years of selection, with two conditions (C1 and C2) applied to the genetic evaluation. In C1 the evaluation considered all the animals until 20th year of selection and in C2 only the animals that were included on the data set used to make the estimation of the genetic trend were considered. The genetic trend was obtained by the adjustment of regression equations, alternatively, through the methods of Ordinary Least Squares (OLS), Weighted Least Squares (WLS) and Generalized Least Squares (GLS), this last one under two forms, one considering homogeneity of variance and another considering heterogeneity of variance throughout the selection years. A quadratic behavior of the genetic values as a function of the year of birth of the animals was verified for the 10 herds studied in the first condition and for herds 1, 4, 8 and 9 in the second condition. For C1 the WLS and GLS methods, although allowing greater inclusion of information in the model, were not able to detect the genetic changes with great precision, the OLS being the method that allowed a greater approach to the annual averages of the real and predicted genetic values. For C2 the GLS method, considering

heterogeneity of variance throughout the years of birth of the animals, was closer to the real genetic changes.

Key words: variance components, multiple regression, heterogeneity of variance, Bayesian inference

I – INTRODUÇÃO GERAL E REVISÃO BIBLIOGRÁFICA

Programas de melhoramento genético são fundamentais para alcançar melhorias na genética dos rebanhos. Porém, depois do programa estar implantado, é necessário que seja feita, periodicamente, a verificação e acompanhamento de sua eficiência (Mello, 1999). Uma das maneiras de realizar um acompanhamento é por meio do conhecimento de parâmetros genéticos e de estimativas de mudança genética, imprescindíveis para o estabelecimento de diretrizes que guiem os programas de melhoramento, para avaliar o progresso genético ao longo do tempo a fim de que os resultados sirvam de elementos orientadores para ações futuras (Euclides Filho et al., 1997).

A tendência genética é uma medida que permite avaliar a mudança ocasionada por um processo de seleção, tornando possível a quantificação da porção genética responsável pelas mudanças acumuladas ao longo dos anos, em determinado rebanho. A estimativa da tendência genética é, até então, a melhor maneira de se observar o progresso genético, visto que a melhora no desempenho não significa obrigatoriamente melhoria genética (Zollinger & Nielsen, 1984; Euclides Filho et al., 1997).

Vários autores utilizam essa metodologia para estudar o comportamento do valor genético ao longo dos anos, dentre eles: Euclides Filho et al. (1997) em rebanhos

Guzerá, Ferraz Filho et al. (1997) em animais da raça Nelore Mocha, Fernandes et al. (2002) na raça Charolês, Silva et al. (1997) e Holanda et al. (2004) em rebanhos Nelore.

A tendência genética é obtida pela regressão dos valores genéticos ao longo do tempo e os coeficientes da regressão são estimados por meio do método de quadrados mínimos, que é baseado na obtenção de estimativas que minimizem a variância residual.

De acordo com Martins et al. (1997) e Oliveira (1998), o método dos quadrados mínimos é dito ordinário (Ordinary Least Square - OLS) quando se considera que os resíduos não são correlacionados e há homocedasticidade. Em casos em que existe heterocedasticidade, porém com resíduos não-correlacionados, o método é dito de quadrados mínimos ponderados (Weighted Least Square - WLS). Se existe correlação entre os resíduos, havendo ou não homogeneidade de variância, pode-se estimar os parâmetros pelo método dos quadrados mínimos generalizados (Generalized Least Square - GLS).

Tendo em vista o modelo linear $y = X\beta + e$, em que y é o vetor de observações, X é a matriz de incidência dos efeitos fixos, β é o vetor dos efeitos fixos e e é o vetor de erros aleatórios, e conforme Searle (1971), os estimadores de quadrados mínimos ordinários (OLS) são obtidos na forma: $\beta = (X'X)^{-1}X'y$; o procedimento de quadrados mínimos ponderados (WLS) é um caso especial de quadrados mínimos generalizados (GLS), sendo estes obtidos por: $\beta = (X'V^{-1}X)^{-1}X'V^{-1}y$, em que a definição para a utilização da matriz V no melhoramento genético e no caso de WLS é $V = I\sigma_e^2$; e no GLS, V é uma matriz não-diagonal, representada por $V = ZGZ' + R$, sendo Z a matriz de incidência dos efeitos aleatórios, G a matriz de (co)variância dos efeitos aleatórios, R a matriz de (co)variância residual; sendo $G = A\sigma_a^2$, em que A é a matriz de correlação entre os efeitos aleatórios nas observações e σ_a^2 é o componente de variância dos efeitos

aleatórios que compõem as observações; $R=I$ $\sigma_{e_i}^2$ é uma matriz identidade de ordem n , e $\sigma_{e_i}^2$ é o componente de variância residual da característica que compõe as observações.

No caso da utilização de GLS, sabe-se que esse método considera a presença dos efeitos aleatórios para a obtenção de estimadores dos efeitos fixos, pois tais efeitos são considerados no modelo, através da matriz de variância e covariância dos efeitos aleatórios que está incluída na matriz V , cuja inversa é utilizada na obtenção dos estimadores (Martins, 1994), e, os estimadores GLS são BLUE (melhor estimador linear não-viciados) e possuem variância mínima, já os estimadores OLS somente serão BLUE quando $V=I\sigma_e^2$, e caso essa exigência não seja cumprida por causa da heterogeneidade, os estimadores permanecem não-viciados, porém não são de variância mínima, sendo viciadas as estimativas das variâncias dos parâmetros estimados.

Os componentes de variância e covariância, imprescindíveis na predição do mérito genético dos indivíduos e utilizados na implementação dos métodos WLS e GLS, têm sido estimados por métodos distintos, de acordo com a evolução de novas teorias e técnicas computacionais. O método da Máxima Verossimilhança Restrita (REML) tem sido muito utilizado na estimação dos componentes, produzindo estimativas pontuais. No entanto, a metodologia Bayesiana vem ganhando espaço na estimação por permitir análise descritiva completa para cada parâmetro, além de permitir a construção de intervalos de credibilidade que levam em conta a incerteza sobre todos os parâmetros do modelo, e por fazer uso de métodos de simulação de Monte Carlo.

De acordo com Nogueira et al. (2003), a metodologia bayesiana é uma ferramenta de grande potencial, pois leva em conta a incerteza existente sobre todos os parâmetros do modelo, além de possibilitar a inclusão de informações passadas, pelo uso de distribuições *a priori* informativas, uma vez que, combinando a verossimilhança com as

prioris, pelo teorema de Bayes, obtém-se a distribuição *a posteriori* conjunta de todos os parâmetros simultaneamente.

Van Tassel & Van Vleck (1996) desenvolveram o software MTGSAM (*Multiple Trait Gibbs Sampling in Animal Model*), que é um conjunto de programas em linguagem Fortran, que utiliza o processo de amostragem de Gibbs em modelos animais, para a estimação bayesiana das médias e distribuições posteriores dos componentes de (co)variância, efeitos fixos e aleatórios e contrastes, em modelo uni e multicarácter, que podem ser utilizado com dados perdidos, com efeitos aleatórios adicionais, correlacionados (efeito genético materno) e não-correlacionados (efeito de ambiente permanente). Além disso, permite a especificação dos efeitos fixos e covariáveis separadamente para cada característica.

O amostrador de Gibbs faz parte de um conjunto de processos iterativos referentes aos métodos de Monte Carlo, baseado em cadeias de Markov (Markov Chain Monte Carlo), os quais tornam aplicáveis à inferência bayesiana em melhoramento genético animal (Magnabosco, 1997).

O uso da amostragem de Gibbs não requer soluções para as equações de modelos mistos de Henderson, permite a análise de conjuntos de dados maiores que os permitidos na metodologia REML, mesmo utilizando rotinas para a resolução de matrizes esparsas, propicia estimativas diretas e acuradas dos componentes de (co)variância e valores genéticos, e pode ser usado em microcomputadores e estações de trabalho, em virtude da pequena quantidade de informação que mantém na memória RAM (Random access memory). Tais características dão certa vantagem computacional ao método de amostragem de Gibbs em relação ao método REML (Van Tassel & Van Vleck, 1996).

O software MTGSAM permite, ainda, a implementação da abordagem que contempla a heterogeneidade de variância, usando modelos multicaracterística em que a característica que se pretende avaliar é tomada como diferentes características em cada classe de heterogeneidade de variância. Assim, é considerado que um indivíduo só possui observação para a característica referente à classe a que ele pertence, apresentando dados perdidos nas características referentes às demais classes. É necessário, no entanto, que haja suficiente conexão genética entre as classes de heterogeneidade para que as estimativas de correlação genética entre as classes sejam precisas, para garantir a confiabilidade das predições dos valores genéticos para uma eficiente seleção dos animais.

Sabe-se que o processo de seleção promove alteração das frequências gênicas e, por consequência, também altera os parâmetros genéticos da população, o que ao longo do tempo pode provocar heterogeneidade de variância e covariância. A não-consideração da heterocedasticidade em diferentes rebanhos, de diferentes regiões, diferentes níveis de manejo e produção e com graus de sangue variados, pode levar a um processo de avaliação genética e seleção viesados, que poderá refletir na escolha de animais que gerarão menor ganho genético, quando utilizados em programas de melhoramento genético (Winkelman & Schaeffer, 1988).

Oliveira (1999), trabalhando com dados de peso aos 365 e 550 dias de idade dos animais pertencentes aos grupos genéticos $1/2$ Charolês + $1/2$ Zebu, $3/4$ Zebu + $1/4$ Charolês, $5/8$ Charolês + $3/8$ Zebu, oriundos do processo de formação da raça Canchim, encontrou diferentes variâncias genéticas e residuais para os diferentes grupos e sugeriu que possivelmente a presença de heterogeneidade de variância teria impacto sobre a avaliação genética dos animais atuais da raça Canchim. Engler (2002), também

trabalhando com animais da raça Canchim, encontrou (co)variâncias heterogêneas entre gerações para peso aos 365 e 550 dias de idade.

A abordagem da avaliação genética, pressupondo heterogeneidade de variância e usando métodos que a contemplem, permite a estimação dos efeitos fixos de ambiente e predição dos valores genéticos, ponderando adequadamente as observações de acordo com as diferentes variâncias em cada classe de heterogeneidade e, ainda, considerando a estrutura de (co)variância entre si (Martins, 2002).

Segundo Martins (2002), quando existe heterogeneidade de variância genética, a fonte de heterocedasticidade afeta a magnitude das diferenças entre os genótipos, mas não afeta a magnitude dos resíduos.

Em programas de melhoramento, tem sido prática nas avaliações genéticas admitir que as variâncias se mantenham constantes ao longo das gerações de seleção, todavia, em rebanhos fechados espera-se que a seleção, além das médias, altere a variância genética aditiva. A dificuldade de se admitir heterocedasticidade ao longo das gerações reside na obtenção de estimativas precisas dos componentes de (co)variância, visto que o número de observações em cada classe de heterocedasticidade se reduz com o aumento do número de classes. Se o número de classes for muito grande, o esforço computacional dispendido para estimar os componentes de (co)variância torna-se muito grande, além do fato de que sendo as classes menos numerosas, as conexões genéticas são mais fracas, levando a estimativas de componentes pouco precisas.

Rekaya et al. (2001) propuseram o uso de um modelo estrutural, em procedimentos bayesianos, como forma de amenizar o problema da estimação de componentes de (co)variância, estimando apenas três parâmetros para acessar 78 componentes de (co)variância. Martins et al. (2002) usaram, com sucesso, um modelo

de regressão polinomial para descrever a trajetória dos componentes de variância ao longo do processo de seleção em gado de corte.

De posse de estimativas precisas das (co)variâncias genéticas heterogêneas, é possível a implementação de métodos para a estimação da tendência genética que busquem melhor avaliação do comportamento do valor genético ao longo das gerações de seleção.

Oliveira (2006), ao trabalhar com tendências genéticas diretas e maternas de características de desempenho em rebanhos de corte simulados, encontrou efeito quadrático do ano de nascimento do animal sobre o valor genético, com incremento dos valores genéticos e redução dos ganhos genéticos ao longo dos anos de seleção.

Alguns autores (Ferraz Filho et al., 1997, 2002; Fernandes et al., 2002; Holanda et al., 2004), ao utilizarem o método de quadrados mínimos ordinários e fazendo o ajuste linear dos valores genéticos em função do ano de nascimento do animal, encontraram taxas de mudança genética anual aquém do potencial sugerido por Smith (1985), que é de 1 a 3% da média da população. Todavia, o método de quadrados mínimos ordinários apresenta propriedades ótimas somente quando os resíduos são não-correlacionados, a distribuição residual é normal e as variâncias são homogêneas, o que nem sempre é a realidade, em função do parentesco entre os animais e do impacto da seleção sobre os parâmetros genéticos.

Em vista disso, o objetivo deste trabalho foi avaliar métodos para a estimação de tendências genéticas, comparando métodos usuais e métodos que contemplem a estrutura de (co)variância entre os valores genéticos ao longo das gerações e o parentesco existente entre os animais.

II – OBJETIVOS GERAIS

Estudar a possibilidade do uso de regressão múltipla para estimar componentes de (co)variância genética aditiva ao longo dos anos pertencentes aos ciclos de seleção em um rebanho fechado, com superposição de gerações.

Avaliar métodos para a estimação de tendências genéticas, comparando métodos usuais e métodos que contemplem a estrutura de (co)variância entre os valores genéticos ao longo das gerações e o parentesco existente entre os animais.

Citações Bibliográficas

- ENGLER, E.O. **Heterogeneidade de variâncias entre gerações para peso aos 365 e 550 dias de idade para animais da raça Canchim.** 2002. 51p. Dissertação (Mestrado em Zootecnia) – Universidade Estadual de Maringá, 2002.
- EUCLIDES FILHO, K.; NOBRE, P.R.C.; ROSA, A.N. Tendência genética na raça Guzerá. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 34., 1997, Juiz de Fora. **Anais...**Juiz de Fora: Sociedade Brasileira de Zootecnia, 1997. p. 175.
- FERNANDES, H.D.; FERREIRA G.B.B.; RORATO, P.R.N. Tendências e parâmetros genéticos para características pré-desmama em bovinos da raça Charolês criados no Rio Grande do Sul. **Revista Brasileira de Zootecnia**, v.31, n.1, p.321-330, 2002.
- FERRAZ FILHO, P.B.; BIANCHINI SOBRINHO, E.; SILVA, L.O.C. et al. Tendência genética em pesos de bovinos da raça Nelore mocha no Brasil. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 34, 1997, Juiz de Fora. **Anais...**Juiz de Fora: Sociedade Brasileira de Zootecnia, 1997.
- FERRAZ FILHO, P.B.; RAMOS, A.A.; SILVA, L.O.C. et al. Tendência genética dos efeitos direto e materno sobre os pesos à desmama e pós-desmama de bovinos da raça Tabapuã no Brasil. **Revista Brasileira de Zootecnia**, v.31, n.2, p.635-640, 2002.
- HOLANDA, M.C.R.; BARBOSA, S.B.P.; RIBEIRO, A.C. et al. Tendências genéticas para crescimento em bovinos Nelore em Pernambuco, Brasil. **Arch. Zootec.**, v.53, p.185-194, 2004.
- 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 de gibbs.** 1997. 83p. Tese (Doutorado em Ciências) – Universidade de São Paulo. Ribeirão Preto, SP, 1997.
- MARTINS, E.N. Uso de modelos mistos no melhoramento animal. In: SIMPÓSIO INTERNACIONAL DE PRODUÇÃO DE NÃO-RUMINANTES, REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 31, Maringá. **Anais...**Maringá: EDUEM, 1994
- MARTINS, E.N. Uso de modelos mistos na avaliação genética animal. Viçosa: UFV, 121p. 1997.

- MARTINS, E.N. Avaliação Genética e Heterogeneidade de Variância. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 39, 2002, Recife-PE. Anais... Recife: Sociedade Brasileira de Zootecnia, [2002]. CD-ROM. Melhoramento genético animal.
- MARTINS, E.N.; ENGLER, E.O.; SAKAGUTI, E.S. et al. In: PROCEEDINGS OF THE 7TH WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7, 2002, Montpellier. **Anais...Montpellier**, v.32, p.303-305, 2002.
- MELLO, S.P. **Tendência genética para pesos em um rebanho da raça Canchim**. Jaboticabal: Universidade Estadual Paulista, 1999. 78p. Dissertação (Mestrado em Zootecnia) – Universidade Estadual Paulista, 1999.
- NOGUEIRA, D.A.; SÁFADI, T.; BEARZOTI, E. et al. Análise clássica e bayesiana de um modelo misto aplicado ao melhoramento animal: uma ilustração. **Ciênc. Agrotec. Lavras**. Edição Especial, p.1614-1624, dez., 2003.
- OLIVEIRA, C.A.L. **Estimadores e preditores em um modelo linear misto, utilizando-se mínimos quadrados ordinários e mínimos quadrados generalizados e a melhor predição linear não viesada**. Maringá: Universidade Estadual de Maringá, 1998. 21p. Monografia (Especialização em Estatística) – Universidade Estadual de Maringá, 1998.
- OLIVEIRA, C.A.L. **Heterogeneidade de variâncias nos grupos genéticos formadores da raça Canchim**. 1999. 64p. Dissertação (Mestrado em Zootecnia) – Universidade Estadual de Maringá, 1999.
- OLIVEIRA, C.A.L. **Efeito direto e materno em características de crescimento em rebanhos de bovinos de corte sob seleção**. 2006. 85p. Tese (Doutorado em Zootecnia) – Universidade Estadual de Maringá, 2006.
- REKAYA, R.; WEIGEL, K.A.; GIANOLA, D. Application of a structural model for genetic covariances in international dairy sire evaluations. **Journal of Dairy Science**, v. 84, p.1525-1530, 2001.
- SEARLE, S.R. **Linear models**. Ithaca: John Wiley & Sons, Inc., 1971. 532p.
- SILVA, L.O.C; FILHO, K.E.; NOBRE, P.R.C. et al. Tendências genéticas na raça Nelore no Brasil. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 34, 1997, Juiz de Fora. **Anais...Juiz de Fora: Sociedade Brasileira de Zootecnia**, 1997a. p.175-176.
- SMITH, C. Rates of genetic change in farm livestock. **Research Development Agricultural**, v.1, n.2, p.79-85; **Animal Breeding Abstracts**, v.53, n.9, p.699, 1985.
- VAN TASSEL, C.P.; VAN VLECK, L.D. Multiple trait Gibbs Sampler for animal models: flexible programs for Bayesian and likelihood-based (co)variance components inference. **Journal of Animal Science**, v.74, p.2586-2597, 1996.
- WINKELMAN, A.; SCHAEFFER, L. R. Effect of heterogeneity of variance on dairy sire evaluation. **Journal of Dairy Science**, v. 71, p.3033-3039, 1988.
- ZOLLINGER, W.A.; NIELSEN, M.K. An evaluation of bias in estimated breeding values for weaning weight in Angus beef cattle field records. I. Estimates of within herd genetic trend. **Journal of Animal Science**, v.58, p.545-549, 1984.

III – Estimação de (co)variâncias genéticas heterogêneas em rebanhos fechados sob seleção

RESUMO – Ao assumir que a seleção em rebanhos fechados pode promover a redução da variância genética aditiva, foi estudada a possibilidade do uso de um modelo de regressão múltipla para estimar os componentes de (co)variância genética aditiva, ao longo dos anos em que a seleção foi praticada. Para tanto, foram usados dados simulados de peso aos 550 dias em 10 rebanhos de bovinos de corte submetidos à seleção por 20 anos. A cada 5 anos, o peso aos 550 dias era uma nova característica, por meio de análises multivariadas envolvendo 4 características, usando Amostrador de Gibbs, com um modelo animal, foram estimados componentes de (co)variância. A convergência das cadeias de Gibbs foi testada pelos métodos de Heidelberg e Welch e de Geweke. Em cada amostra gerada, formada por 10 componentes de (co)variância genética aditiva, foi ajustada uma equação de regressão múltipla para estimar os 406 componentes de (co)variância genética aditiva durante os 20 anos de seleção. Os coeficientes de determinação (R^2) da regressão múltipla apresentaram-se altos e pouco variáveis, com média 0,9027 e desvio-padrão de 0,0394. Observou-se que o modelo de regressão múltipla pode ser usado na estimação dos componentes de (co)variância genética quando se admite heterocedasticidade ao longo do tempo, causada pela seleção, podendo ser implementado por procedimentos bayesianos.

Palavras-chave: componentes de (co)variância, heterocedasticidade, inferência Bayesiana, regressão múltipla

III – Estimation of heterogeneous genetic (co)variances in closed herds under selection

ABSTRACT – Assuming that the selection in closed herds can promote the reduction of the additive genetic variance, the possibility of the use of a multiple regression model was studied to estimate the additive genetic (co)variance components, throughout the years when the selection was carried out. For that, weights at 550 days (P550) were studied using simulated data of herds submitted to 20 years of selection. The (co)variance components were estimated assuming that each 5 years the weight at 550 days was a new characteristic, by means of the multi-character analyses involving 4 characteristics, using the Gibbs sampler in an animal model. The convergence of the Gibbs chains was tested by the Heidelberg and Welch and the Geweke methods. In each sample generated, formed for 10 components of additive genetic (co)variance, an equation of multiple regression was adjusted to estimate the 406 additive genetic (co)variance components during the 20 years of selection. The determination coefficients (R^2) of the multiple regression were high and steady, with mean of 0.9027 and standard deviation of 0.0394. The results indicate that the multiple regression model of (co)variance components as a function of generations under selection can be introduced in Bayesian procedures for the estimation of (co)variance components, when heteroscedasticity due to selection is assumed along time.

Key words: variance components, heteroscedasticity, Bayesian inference, multiple regression

Introdução

A abordagem da avaliação genética, ao pressupor heterogeneidade de variância e usando métodos que a contemple, permite a estimação dos efeitos fixos de ambiente e predição dos valores genéticos, ponderando adequadamente as observações de acordo com as diferentes variâncias em cada classe de heterogeneidade e, ainda, considerando a estrutura de (co)variância entre si (Martins, 2002).

Em programas de melhoramento, tem sido prática nas avaliações genéticas admitir que as variâncias se mantenham constantes ao longo das gerações de seleção, todavia, em rebanhos fechados espera-se que a seleção, além das médias, altere a variância genética aditiva. A dificuldade de se admitir heterocedasticidade ao longo das gerações reside na obtenção de estimativas precisas dos componentes de (co)variância, visto que o número de observações em cada classe de heterocedasticidade se reduz com o aumento do número de classes. Se o número de classes for muito grande, o esforço computacional dispendido para estimar os componentes de (co)variância torna-se muito grande, além do fato de que sendo as classes menos numerosas, as conexões genéticas são mais fracas, levando a estimativas de componentes pouco precisas.

Rekaya et al. (2001) propuseram o uso de um modelo estrutural, em procedimentos bayesianos, como forma de amenizar o problema da estimação de componentes de (co)variância, estimando apenas três parâmetros para acessar 78 componentes de (co)variância. Martins et al. (2002) usaram, com sucesso, um modelo de regressão polinomial para descrever a trajetória dos componentes de variância ao longo do processo de seleção em gado de corte.

Desta forma, este trabalho teve como objetivo estudar a possibilidade do uso de regressão múltipla para estimar componentes de (co)variância genética aditiva ao longo dos anos de seleção em um rebanho fechado, com superposição de gerações.

Material e Métodos

O conjunto de dados utilizado, formado por rebanhos de bovinos de corte submetidos a 20 anos de seleção, foi simulado em linguagem de programação Fortran, compilador F90.

Nesses rebanhos, o plantel de reprodução era formado de 1.500 vacas e 38 touros, mantendo-se na monta natural uma relação touro-vaca de 1:40. A partir do segundo ano, foi implementada inseminação artificial em 50% das vacas, sendo utilizados 10 touros para inseminação artificial e 19 touros para monta natural. A taxa de natalidade estabelecida foi de 90% e sobrevivência de 95% até o início da reprodução.

A taxa de descarte de vacas era variável e determinada pelo número de vacas vazias ao final da estação-monta. Apenas as vacas primíparas poderiam permanecer no plantel de reprodução se estivessem vazias. Os animais jovens poderiam ser incluídos no plantel de reprodução a partir de 22 meses de idade.

Foram simulados dados de peso ao desmame (PD) e aos 550 dias (P550), assim como os valores genéticos para efeito direto de peso ao desmame e peso aos 550 dias e efeito materno de peso ao desmame.

A cada ano procedia-se a avaliação genética com o objetivo de orientar a seleção, utilizando-se as equações de modelos mistos (Henderson, 1984) para predição dos valores genéticos, por meio de análise bicaracter, no sistema computacional MTDFREML (*Multiple Trait Derivative Free Restricted Maximum Likelihood*) (Boldman et al. 1995). Os efeitos identificáveis de ambiente foram: sexo (macho ou fêmea), época de nascimento (início, meio ou final da estação de nascimento), ano de nascimento e idade (em meses) da mãe ao parto. Os valores dos componentes de (co)variância genética inicialmente utilizados foram os seguintes,

$$G_0 = \begin{bmatrix} \sigma_{aD}^2 & \sigma_{aDm} & \sigma_{aDS} \\ \sigma_{aDm} & \sigma_m^2 & \sigma_{aSm} \\ \sigma_{aDS} & \sigma_{aSm} & \sigma_{aS}^2 \end{bmatrix} = \begin{bmatrix} 165 & -55 & 135 \\ -55 & 95 & 0 \\ 135 & 0 & 215 \end{bmatrix}$$

em que:

$\sigma_{a_i}^2$ é a variância genética aditiva para característica i , $i=PD(D)$ e $P550(S)$;

σ_m^2 é a variância genética aditiva materna da característica PD;

$\sigma_{a_{DS}}$ é a covariância genética entre os efeitos diretos de PD e P550;

$\sigma_{a_{i,m}}$ é a covariância genética entre os efeitos diretos das características i com os

efeitos maternos de PD, $i=PD(D)$, $P550(S)$;

As variâncias residuais adotadas foram 365 e 450 kg² para PD e P550, respectivamente, e o valor da covariância residual foi de 150 kg².

Para classificação dos candidatos à seleção, utilizou-se um índice empírico de seleção com pesos de 0,3; 0,3; 0,4, respectivamente, para os valores genéticos direto e materno para PD e direto para P550.

Os acasalamentos dos animais selecionados foram realizados ao acaso, com restrição para controlar o incremento dos níveis de endogamia, por meio de impedimento do acasalamento entre pais e filhos e entre irmãos (meio-irmãos ou irmãos completos).

Após o estabelecimento das condições de simulação e seleção, foram realizadas 10 repetições, totalizando 10 rebanhos com média de 31.198 animais, como descrito por Oliveira (2006).

Para a obtenção das (co)variâncias genéticas entre os anos de nascimento, para a característica peso aos 550 dias, primeiramente foi realizada a estimação dos componentes de variância de P550 de cada geração de seleção (5 anos) e covariância de

P550 entre as gerações, por meio do software MTGSAM (*Multiple Trait Gibbs Sampling in Animal Model*), em análise multicarácter em que o peso aos 550 dias foi tratado como 4 características distintas de acordo com a geração em que o animal nasceu, sendo P550-g1 o peso aos 550 dias para os animais nascidos do ano 1 ao ano 5, P550-g2 para os nascidos entre 6 e 10, P550-g3 para 11 ao 15 e P550-g4 para os nascidos entre os anos 16 e 20. Foi considerado que um animal só possui observação para a característica referente à classe a que ele pertence, apresentando dados perdidos nas características referentes às demais classes. Sexo, estação de nascimento e ano de nascimento foram considerados efeitos fixos, e idade da mãe ao parto foi considerada como covariável.

Embora o MTGSAM seja um programa flexível quanto aos modelos de análise que podem ser utilizados, as distribuições *a priori* dos efeitos considerados no modelo são definidas pelo programa, sendo necessário definir apenas alguns parâmetros da distribuição dos componentes de variância (Lôbo et al., 1997).

Assim, foi assumido que para os efeitos de ambiente identificáveis não existe nenhum conhecimento inicial, tendo uma distribuição inicial uniforme e que a distribuição dos efeitos genéticos e dos resíduos é normal multivariada. Em se tratando dos valores genéticos, foi considerada a estrutura de (co)variância conhecida pela matriz de parentesco. As definições para os elementos do modelo animal e da distribuição conjunta de Y , a e e são, respectivamente:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} X_1 & 0 & 0 & 0 \\ 0 & X_2 & 0 & 0 \\ 0 & 0 & X_3 & 0 \\ 0 & 0 & 0 & X_4 \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} + \begin{bmatrix} Z_1 & 0 & 0 & 0 \\ 0 & Z_2 & 0 & 0 \\ 0 & 0 & Z_3 & 0 \\ 0 & 0 & 0 & Z_4 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix};$$

em que:

y_1, y_2, y_3, y_4 são os vetores de observações, para peso aos 550 dias de idade, referentes às gerações 1, 2, 3 e 4, respectivamente;

X_1, X_2, X_3, X_4 são as matrizes de incidência dos efeitos identificáveis de ambiente, para peso aos 550 dias de idade, referentes às gerações 1, 2, 3 e 4, respectivamente;

$\beta_1, \beta_2, \beta_3, \beta_4$ são os vetores dos efeitos identificáveis de ambiente, para peso aos 550 dias de idade, referentes às gerações 1, 2, 3 e 4, respectivamente;

Z_1, Z_2, Z_3, Z_4 são as matrizes de incidência dos efeitos aleatórios, peso aos 550 dias de idade, referentes às gerações 1, 2, 3 e 4, respectivamente;

a_1, a_2, a_3, a_4 são os vetores dos efeitos genéticos diretos, para peso aos 550 dias de idade, referentes às gerações 1, 2, 3 e 4, respectivamente;

e_1, e_2, e_3, e_4 são os vetores dos erros aleatórios, para peso aos 550 dias de idade, referentes às gerações 1, 2, 3 e 4, respectivamente; com distribuição conjunta:

$$\begin{bmatrix} y \\ a \\ e \end{bmatrix} \sim NMV \left\{ \begin{bmatrix} X\beta \\ 0 \\ 0 \end{bmatrix}; \begin{bmatrix} ZGZ'+R & ZG & R \\ GZ' & G & 0 \\ R & 0 & R \end{bmatrix} \right\}, \text{ em que:}$$

$$G = G_0 \otimes A$$

\otimes é o produto de Kronecker;

G_0 é a matriz de (co)variância genética das i gerações, para peso aos 550 dias de idade, dada a seguir:

$$G_0 = \begin{bmatrix} \sigma_{a_1}^2 & \sigma_{a_1 a_2} & \sigma_{a_1 a_3} & \sigma_{a_1 a_4} \\ \sigma_{a_1 a_2} & \sigma_{a_2}^2 & \sigma_{a_2 a_3} & \sigma_{a_2 a_4} \\ \sigma_{a_3 a_1} & \sigma_{a_3 a_2} & \sigma_{a_3}^2 & \sigma_{a_3 a_4} \\ \sigma_{a_4 a_1} & \sigma_{a_4 a_2} & \sigma_{a_4 a_3} & \sigma_{a_4}^2 \end{bmatrix}, \text{ sendo:}$$

$\sigma_{a_i}^2$ é a variância genética aditiva da geração i ; $i = 1, 2, 3$ e 4

$\sigma_{a_i a_j}$ é a covariância genética entre as gerações i e j ;

$i = 1, 2, 3$ e $j = 1, 2, 3$ e 4

$$R = R_0 \otimes I$$

em que:

I é a matriz identidade de ordem igual ao número de animais;

R_0 é a matriz de variância residual, das i gerações, para peso aos 550 dias de idade, dada a seguir:

$$R_0 = \begin{bmatrix} \sigma_{e_1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{e_2}^2 & \dots & 0 \\ \vdots & \vdots & \sigma_{e_3}^2 & \vdots \\ 0 & 0 & \dots & \sigma_{e_4}^2 \end{bmatrix}, \text{ em que:}$$

$\sigma_{e_i}^2$ é a variância residual da geração i ; $i = 1, 2, 3$ e 4 .

Para os efeitos ambientais identificáveis foi assumida distribuição “uniforme”.

Para os componentes de (co)variância genética foi assumido que G têm distribuição de Wishart Invertida (IW), na forma

$$G \sim IW(G_0, \nu_g),$$

tal que

$$f(G/G_0, \nu_g) \propto |G|^{-\frac{1}{2}(\nu_g + m_g + 1)} \cdot e^{-\frac{1}{2}tr(G_0^{-1}G^{-1})}$$

em que:

ν_g é o grau de liberdade da distribuição que equivale ao grau de crença que se tem acerca do parâmetro,

m_g é a ordem da matriz G ,

G_0 é a matriz de parâmetro escala da distribuição de G .

Para os componentes de variância residual foi assumido que cada elemento $\sigma_{e_i}^2$ de R_0 , tem distribuição qui-quadrado invertida, na forma

$$\sigma_{e_i}^2 \sim IX^2 (S_{0i}, \nu_{r_i})$$

tal que:

$$f(\sigma_{e_i}^2/S_{0i}, \nu_{r_i}) \propto (\sigma_{e_i}^2)^{-\frac{1}{2}(\nu_{r_i}+2)} \cdot e^{-\frac{1}{2}(S_{0i}/\sigma_{e_i}^2)}$$

em que:

ν_{r_i} é o grau de liberdade da distribuição que equivale ao grau de crença que se tem acerca dos parâmetro,

S_{0i} é o parâmetro escala da distribuição de $\sigma_{e_i}^2$.

São calculados como segue:

$$G_0 = \frac{G^*}{\nu_g^*} = \frac{G^*}{(\nu_g - m_g - 1)}$$

em que G^* é a matriz de *priors* dos componentes de variância contidos em G .

$$S_{0i} = \frac{\sigma_{e_i}^{2*}}{\nu_{r_i}^*} = \frac{\sigma_{e_i}^{2*}}{(\nu_{r_i} - 2)},$$

em que $\sigma_{e_i}^{2*}$ é o *prior* do componente de variância residual $\sigma_{e_i}^2$.

A função densidade de probabilidade conjunta posterior dos parâmetros, dados os hiperparâmetros, foi obtida por meio do produto das distribuições *a priori* pela função de verossimilhança, conforme descrito por Van Tassel & Van Vleck (1995). Para implementar a amostragem de Gibbs, foram obtidas as funções densidade de probabilidade condicionais completas para cada parâmetro, por meio da derivação da

função de densidade posterior conjunta, considerando como constantes os valores dos outros parâmetros contidos no modelo, sendo gerada a distribuição do parâmetro de interesse:

$$\begin{aligned}
 & f(\beta, u, G, R \mid y, v_g, G_0, v_{e_1}, s_{0_1}, \dots, v_{e_4}, s_{0_4}) \\
 & \propto |R|^{-1/2} \cdot \exp\left\{-\frac{1}{2}(y - X\beta - Zu)' R^{-1}(y - X\beta - Zu)\right\} \\
 & \quad \cdot |G|^{-\frac{1}{2}(n+v_g+m_g+1)} \cdot \exp\left\{-\frac{1}{2}[a'(G^{-1} \otimes A^{-1})a + \text{tr}(v_g^* G_0 G^{-1})]\right\} \\
 & \quad \cdot \prod_{i=1}^7 \left[(\sigma_{e_i}^2)^{-1/2}(q_i+v_{e_i}+2) \cdot \exp\left\{-\frac{v_{e_i}^*}{2} s_{0_i} \sigma_{e_i}^{-2}\right\} \right].
 \end{aligned}$$

As estimativas das médias *posteriores* para os componentes de (co)variância foram baseadas no valor esperado de suas distribuições condicionais completas.

Para cada uma das repetições (rebanhos) foi gerada uma cadeia de Gibbs de 1.000.000 de ciclos com amostras retiradas a cada 100 ciclos, após a eliminação de 50.000 ciclos iniciais, gerando cadeias de 9500 amostras dos componentes de (co)variância. Foi diagnosticada a convergência das cadeias de amostragem de Gibbs ao utilizar os métodos de Geweke (1992) que é um teste paramétrico simples que utiliza a estatística Z para comparar médias amostrais de dois lugares distintos ao longo da cadeia, e o de Heidelberger & Welch (1983) que em primeira instância compara a cadeia de Gibbs com uma cadeia hipotética de distribuição estacionária, para verificar se as médias amostrais estão dentro de um limiar do intervalo de credibilidade, estabelecido pelo teste. Ambos os testes estão disponíveis na biblioteca CODA (Convergence Diagnosis and Output Analysis, versão 0.4), desenvolvida por Cowles et al. (1995), e implementada no software R (2004).

Para a obtenção das (co)variâncias genéticas para P550 entre e para os anos em que a seleção foi realizada, foi utilizado um modelo de regressão múltipla em que: $Y = b_0 + b_1x_1 + b_2x_2 + b_3x_1^2 + b_4x_2^2 + b_5x_1x_2$, em que:

Y é o componente de (co)variância que se pretende estimar;

b_0 é a constante geral;

b_1, b_2, b_3, b_4 e b_5 são os coeficientes da regressão;

x_1 e x_2 são os anos de nascimento dos animais.

Por meio de programas desenvolvidos em linguagem Fortran, compilador F90, e utilizada a amostragem dos componentes de (co)variância genética para geração de seleção fornecidos pelo software MTGSAM, foi possível a estimação dos coeficientes da regressão múltipla pelo método dos quadrados mínimos generalizados (GLS), conforme:

$$\beta = (X'V^{-1}X)^{-1}X'V^{-1}y, \text{ em que:}$$

β é o vetor dos coeficientes da regressão;

X é a matriz de incidência dos efeitos fixos de ano;

V é a matriz de (co)variância entre os 10 componentes de (co)variância genética entre gerações, estimados pelo software MTGSAM;

y é o vetor de médias posteriores dos componentes estimados pelo software MTGSAM.

O modelo de regressão múltipla produz uma superfície de resposta, que permite, a partir dos pontos existentes, fazer inferência sobre qualquer ponto dentro do limite estudado.

Para cada rebanho, foi ajustada uma equação que permitiu a obtenção da matriz de (co)variância genética de P550 entre os anos de nascimento dos animais, sendo que os componentes de (co)variância de P550 para os animais nascidos anteriormente ao primeiro ano de seleção e do primeiro ao quinto ano de seleção receberam os valores de variância e covariância do primeiro ano de seleção.

Para a observação da dispersão dos componentes de (co)variância e para saber se os valores estimados por meio da regressão eram condizentes com a realidade e se estavam contidos no intervalo estipulado em 90% da probabilidade total, foi calculado o intervalo de credibilidade a 90%, utilizando-se a biblioteca CODA (Convergence Diagnosis and Output Analysis, versão 0.4), desenvolvida por Cowles et al. (1995), e implementado no software R (2004).

Foi realizada a comparação entre os intervalos de credibilidade das variâncias genéticas entre gerações para ver se, de fato, as variâncias poderiam ser consideradas heterogêneas.

O poder de inferência sobre os componentes de variância genética para os anos de nascimento e covariância genética entre os anos de nascimento dos animais, obtidos a partir da regressão múltipla, foi testado pelo coeficiente de determinação, representado pela razão entre a soma de quadrados dos valores de (co)variância estimados pela regressão para as 4 gerações, e a soma de quadrados dos valores de (co)variância fornecidos pelo MTGSAM, para as mesmas 4 gerações.

Resultados e Discussão

Para a convergência das cadeias de Gibbs, o método de Heidelberger & Welch (1983) apontou falha na convergência do componente de covariância genética para P550 dias entre a primeira e a quarta geração de seleção para os rebanhos 1 e 3 e falha na convergência do componente de covariância genética de P550 entre a segunda e a quarta geração para os rebanhos 5 e 8. O método de Geweke (1992) mostrou que a convergência do componente de covariância genética para P550 entre a terceira e quarta geração falhou para o rebanho 4 e o componente de covariância entre a segunda e a

quarta geração falhou para o rebanho 9, ao indicar que estes componentes não atingiram a fase de estacionariedade. Todavia, o fato do teste de Heidelberger e Welch reprovar cadeias aprovadas no teste de Geweke e vice-versa, é explicado pela consideração de critérios distintos na verificação da convergência nos dois testes.

O alcance da convergência depende de alguns fatores, como o intervalo de retirada das amostras e o tamanho da cadeia. Segundo Sorensen (1996), é difícil ter uma garantia empírica positiva de que uma cadeia de Gibbs é longa o suficiente, uma vez que o GS pode executar muitas iterações em uma certa região do espaço amostral, dando a impressão da convergência, e posteriormente, mover-se para outra região. Contudo, Simonelli (2004), ao estudar a heterogeneidade de variância para peso em gado de corte verificou que, mesmo sem alcançar a convergência, as médias posteriores obtidas a partir de cadeias longas foram adequadas para estimar componentes de (co)variância, não sendo porém recomendado inferências quanto às distribuições posteriores.

A dificuldade verificada para os componentes, acima citados, atingirem a convergência, pode ser explicada pela distância existente entre as gerações, que resulta em menor conexão entre elas.

Para os 10 rebanhos estudados, foi observada a diminuição das variâncias genéticas para P550 dias à medida que as gerações de seleção avançaram, e redução das covariâncias genéticas entre as gerações proporcional à distância entre as gerações, o que pode ser verificado na Tabela 1.

Tabela 1 – Componentes de variância genética aditiva estimados por geração para P550
 Table 1 – Components of additive genetic variance estimated for generation, for P550

Componentes <i>Components</i>	Rebanhos <i>Herds</i>				
	1	2	3	4	5
$\sigma_{a_1}^2$	148,699133	182,608436	139,619713	182,277422	159,451658
$\sigma_{a_2}^2$	132,764510	119,560177	127,189238	116,686398	137,023988
$\sigma_{a_3}^2$	66,596612	117,098940	63,083555	81,199387	59,467556
$\sigma_{a_4}^2$	66,009022	91,677928	56,199304	59,728454	48,656739
$\sigma_{a_1a_2}$	77,128814	78,921871	74,772510	77,252624	77,962332
$\sigma_{a_1a_3}$	28,683212	66,497432	19,077383	48,319568	27,029108
$\sigma_{a_1a_4}$	19,629580	26,727634	13,912283	21,472257	15,139163
$\sigma_{a_2a_3}$	26,415454	50,731095	19,031811	42,451648	24,083913
$\sigma_{a_2a_4}$	24,421595	25,157914	17,499574	19,855809	12,298283
$\sigma_{a_3a_4}$	12,019043	29,896309	17,673280	18,373771	8,160788
Componentes <i>Components</i>	6	7	8	9	10
$\sigma_{a_1}^2$	144,403158	168,927513	151,083970	144,954762	168,395903
$\sigma_{a_2}^2$	112,123176	128,119851	116,139435	142,108701	139,612526
$\sigma_{a_3}^2$	78,282904	88,384501	71,953876	91,707890	94,702230
$\sigma_{a_4}^2$	61,867688	84,822042	55,492652	79,798405	70,681478
$\sigma_{a_1a_2}$	65,115592	91,864556	63,877775	98,822019	97,044499
$\sigma_{a_1a_3}$	36,360779	53,093465	22,702495	42,663531	53,180222
$\sigma_{a_1a_4}$	20,210127	31,401905	18,518415	27,084457	23,588824
$\sigma_{a_2a_3}$	25,164168	49,528593	16,544610	47,065025	50,939175
$\sigma_{a_2a_4}$	26,726232	32,608890	10,989712	24,915798	24,909425
$\sigma_{a_3a_4}$	17,559843	22,776102	10,206402	18,515839	13,601500

Para Gomez-Raya & Burnside (1990), quanto maior a acurácia de seleção maior é a redução das variâncias, sendo a acurácia de seleção definida como a correlação genética entre o valor genético verdadeiro e o predito. De acordo com Quinton & Smith (1995) e Martins (2003), o decréscimo acentuado das variâncias genéticas está relacionado com a utilização das informações de parentes na predição dos valores genéticos, por meio da matriz de parentesco, que pode gerar co-seleção, aumentando a

probabilidade de seleção de animais aparentados, e as perdas, na variabilidade genética, estão relacionadas com altos níveis de consangüinidade em populações sob seleção. No presente caso, as perdas foram menos bruscas porque os níveis de consangüinidade foram mantidos baixos pelo impedimento de acasalamentos entre pais e filhos e entre irmãos.

Ferraz Filho et al. (2002) encontraram valor de 225,06 kg² para variância genética aditiva para peso aos 550 dias, em animais da raça Tabapuã. Ao trabalhar com animais da raça Canchim, Pereira (2003) encontrou variância genética de 420,371 kg² para peso ao sobreano. Em animais da raça Nelore, Van Melis et al. (2003) encontraram, para variância genética de peso aos 550 dias, valor de 205,60 kg². Estes valores são superiores aos aqui encontrados, no entanto, os autores citados consideraram homogeneidade de variância ao longo dos anos.

Engler (2002), trabalhando com animais da raça Canchim, encontrou, para a variância genética aditiva aos 365 dias, valores decrescentes da primeira à quinta geração de seleção. Para peso aos 550 dias, os valores decresceram da primeira à terceira geração e aumentaram da quarta à sexta geração. Para as covariâncias genéticas, foram encontrados valores mais elevados para gerações mais próximas à medida que as gerações eram mais afastadas, as covariâncias para peso aos 365 e 550 dias diminuíam, da primeira à sexta e da primeira à quinta gerações, respectivamente.

Os intervalos de credibilidade (IC) para as estimativas de variância genética aditiva para geração de seleção podem ser visualizados na Tabela 2.

Tabela 2 – Intervalos de Credibilidade (IC) para variâncias genéticas de geração de seleção

Table 2 – Credibility intervals for genetic variances of the selection's generation

Rebanhos <i>Herds</i>	Componentes <i>Components</i>							
	$\sigma_{a_1}^2$	$\sigma_{a_2}^2$	$\sigma_{a_3}^2$	$\sigma_{a_4}^2$	$\sigma_{a_1}^2$	$\sigma_{a_2}^2$	$\sigma_{a_3}^2$	$\sigma_{a_4}^2$
1	121,24	-- 178,91	106,48	-- 160,29	49,05	-- 86,32	47,50	-- 87,06
2	154,52	-- 211,47	97,40	-- 143,98	92,38	-- 143,36	67,71	-- 118,71
3	113,84	-- 166,80	102,60	-- 153,88	46,44	-- 83,33	40,00	-- 75,00
4	152,33	-- 214,14	94,53	-- 141,36	61,31	-- 102,65	42,07	-- 80,71
5	133,45	-- 187,71	109,36	-- 166,07	42,77	-- 78,21	33,09	-- 66,41
6	119,98	-- 170,87	88,57	-- 137,89	59,11	-- 99,33	44,47	-- 81,34
7	142,51	-- 196,55	104,73	-- 153,70	68,99	-- 109,96	63,30	-- 109,41
8	125,14	-- 178,32	92,74	-- 141,92	52,21	-- 95,10	38,47	-- 74,39
9	122,27	-- 169,38	119,38	-- 166,67	70,68	-- 114,83	58,69	-- 104,57
10	140,94	-- 196,98	114,93	-- 166,39	72,19	-- 120,87	51,76	-- 92,09

Observou-se que em relação aos rebanhos 1 e 7, os IC foram coincidentes para as variâncias da geração 3 e 4, e no rebanho 2 coincidiram as variâncias para geração 2 e 3, com a indicação de homocedasticidade nestes casos, e que no restante dos rebanhos existe heterogeneidade de variâncias, pois os IC não foram coincidentes. Engler (2002) verificou que a não-consideração da heterogeneidade de (co)variâncias causou impacto na avaliação genética para a característica peso aos 550 dias.

Nuñez-Domingues et al. (1995), em experimento com 12 raças de reprodutores, cruzados com fêmeas Aberdeen Angus e Hereford, encontraram melhor ajuste para o modelo que considerava variâncias heterogêneas e verificaram alterações nas classificações fornecidas pelos modelos homocedásticos e heterocedásticos.

De acordo com Winkelman & Schaeffer (1988), a não-consideração de heterogeneidade de variância em diferentes rebanhos, de diferentes regiões, diferentes níveis de manejo e produção e com graus de sangue variados, pode levar a um processo de avaliação genética e seleção viesados que poderá refletir na escolha de animais que gerarão menor ganho genético, quando utilizados em programas de melhoramento genético.

Oliveira (2006) citando Ehlers (2003) expõe que o tamanho do intervalo de credibilidade está associado à dispersão do parâmetro, de maneira que quanto menor o intervalo mais concentrada é a distribuição a posterior dos parâmetros e que a definição de intervalo de credibilidade expressa, por meio de probabilidade, a pertinência ou não do parâmetro ao intervalo.

Os coeficientes lineares e quadráticos da regressão múltipla estimados para cada rebanho podem ser vistos na Tabela 3. Calculados a partir dos componentes de (co)variância para geração de seleção, eles permitiram um ajuste para cada rebanho, de uma equação para a estimação dos componentes heterogêneos de (co)variância genética por ano de nascimento do animal.

Tabela 3 – Valores dos coeficientes da regressão estimados para cada rebanho e coeficientes de determinação (R^2) das equações ajustadas

Table 3 – Values of regression coefficients estimated for each herd and determination coefficients (R^2) of the adjusted equations

Rebanhos							
<i>Herds</i>	b_0	b_1	b_2	b_3	b_4	b_5	R^2
1	211,55864	13,71237	-25,57581	0,39999	0,88612	-1,08491	0,874863
2	240,31653	3,20521	-19,32153	0,73711	0,59534	-0,88641	0,838414
3	194,10682	14,24208	-24,74680	0,34322	0,85045	-1,02171	0,871140
4	255,54433	5,80474	-23,99069	0,44112	0,71401	-0,73665	0,906774
5	231,17821	12,27233	-26,25605	0,42355	0,86808	-1,06398	0,919815
6	196,12719	10,89269	-22,41507	0,41593	0,78600	-0,97032	0,866785
7	230,71259	5,67336	-19,81298	0,53353	0,59789	-0,79931	0,915144
8	214,58662	13,34926	-26,60404	0,54575	0,95587	-1,23993	0,960663
9	157,69636	11,08293	-12,34808	0,55181	0,38810	-1,08897	0,919413
10	201,88020	9,48369	-16,11425	0,48423	0,45404	-0,95110	0,954173

b_0 é a constante geral; b_1, b_2, b_5 são coeficientes lineares; b_3, b_4 são coeficientes quadráticos

b_0 is the general constant; b_1, b_2, b_5 are linear coefficients; b_3, b_4 are quadratic coefficients

Quando foi testado o grau de ajuste e o poder de inferência da regressão múltipla para a estimação dos componentes para ano, a partir dos componentes para geração, foram obtidos altos coeficientes de determinação (R^2), para todos os rebanhos, com média de 0,9027 e desvio-padrão de 0,0394. Os coeficientes de determinação podem ser vistos na Tabela 3.

Partindo dos 10 componentes de (co)variância para as 4 gerações de seleção, foi possível a estimação de 28 componentes de variância e 378 componentes de covariância, que totalizou 406 componentes. Os valores das variâncias genéticas para cada ano de nascimento dos animais podem ser vistos na Tabela 1A do Apêndice. Para os anos 1 ao 13, foi considerada a mesma variância porque a seleção foi iniciada a partir do 9º ano, e deste até o 13º foi considerada a primeira geração de seleção. Os componentes de covariância podem ser vistos nas Tabelas 2A e 3A do Apêndice. As covariâncias entre os anos, do ano 1 ao 13 foram consideradas iguais pelo mesmo motivo acima citado.

As variâncias genéticas estimadas para ano de nascimento não apresentaram o mesmo comportamento das variâncias genéticas por geração de seleção, que foram reduzidas à medida que as gerações se passaram, para todos os rebanhos. Foi verificada a redução das variâncias, por ano, por completo, ou seja, do 1º ao 28º ano, para os rebanhos 1, 3, 5, 9 e 10. Para o rebanho 2, foi diminuída do ano 1 ao 18, e aumentada do 19 ao 28. No rebanho 4, a redução foi do ano 1 ao 23, elevando-se do 24 ao 28. Nos rebanhos 6 e 8 foi reduzida do 1º ao 25º ano e aumentada do 26º ao 28º. Para o rebanho 7, o decréscimo foi do 1º ao 21º ano, aumentando do 22º ao 28º. Esse comportamento de redução e posterior acréscimo das variâncias, observado em alguns rebanhos, pode ser explicado pela utilização de um critério de acasalamento de animais não-aparentados, o uso de diferentes técnicas de reprodução, taxa de descarte dos animais, que pode ter deixado de favorecer determinados reprodutores e pela taxa de reposição.

De maneira geral, foi constatada a redução das covariâncias genéticas entre os anos dentro da mesma geração, à medida que os anos se distanciaram; fato que pode ser explicado pela diminuição da conexão entre os animais nascidos em anos mais distanciados. A redução das covariâncias, uma vez que foi observada a redução nas

variâncias, é também um indício de que, segundo Gianola et al. (1992), quando há heterogeneidade de variância, certamente também deve haver heterogeneidade de covariância.

O comportamento da superfície de resposta produzida pelo modelo de regressão múltipla pode ser visualizado na Figura 1, para o rebanho 1.

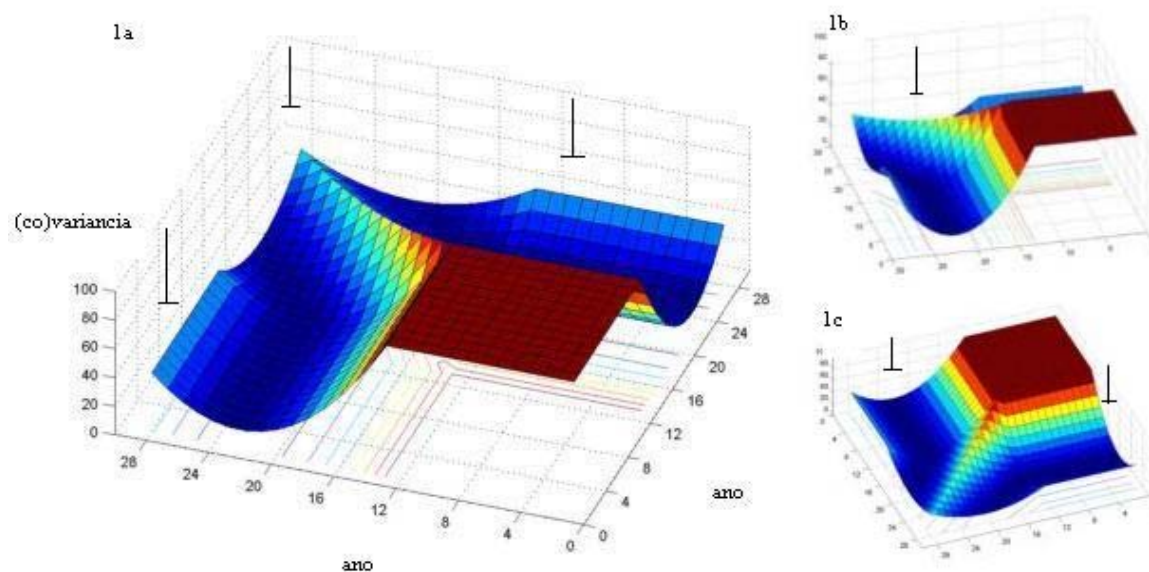


Figura 1 – Superfície de resposta para (co)variâncias genéticas por ano de nascimento, para P550, para o rebanho 1. O gradiente vermelho até azul é traduzido por valores maiores até valores menores. 1a mostra a redução das variâncias e covariâncias, 1b mostra a redução das variâncias e em 1c a redução das covariâncias.

Figure 2 – Response surface of genetic (co)variances for year of birth, for P550, for the first herd. The red-to-blue gradient is translated by greater values until lower values. 1a shows the variance and covariance reduction, 1b shows the variance reduction, and 1c the covariance reduction.

A Figura 1a permite a visualização do decréscimo tanto das variâncias como das covariâncias, conforme indicado pelas setas, a 1b mostra claramente a redução das variâncias e na Figura 1c está indicada a redução das covariâncias.

Comparando-se o comportamento das variâncias genéticas ao longo das gerações com as variâncias ao longo dos anos, pôde-se observar que as últimas mostraram-se subestimadas em relação às primeiras, o que pode ser visto nas Figuras 2 e 3.

Quando foi verificado se os intervalos de credibilidade para variância genética de geração de seleção continham as estimativas das variâncias genéticas para ano de nascimento do animal, observou-se que a maioria dos componentes encontravam-se fora dos ICs, sendo apenas 17,40%, em média, das estimativas contidas nos respectivos ICs; isso fornece indícios de que as estimativas para o ano podem não ser condizentes com seus valores reais.

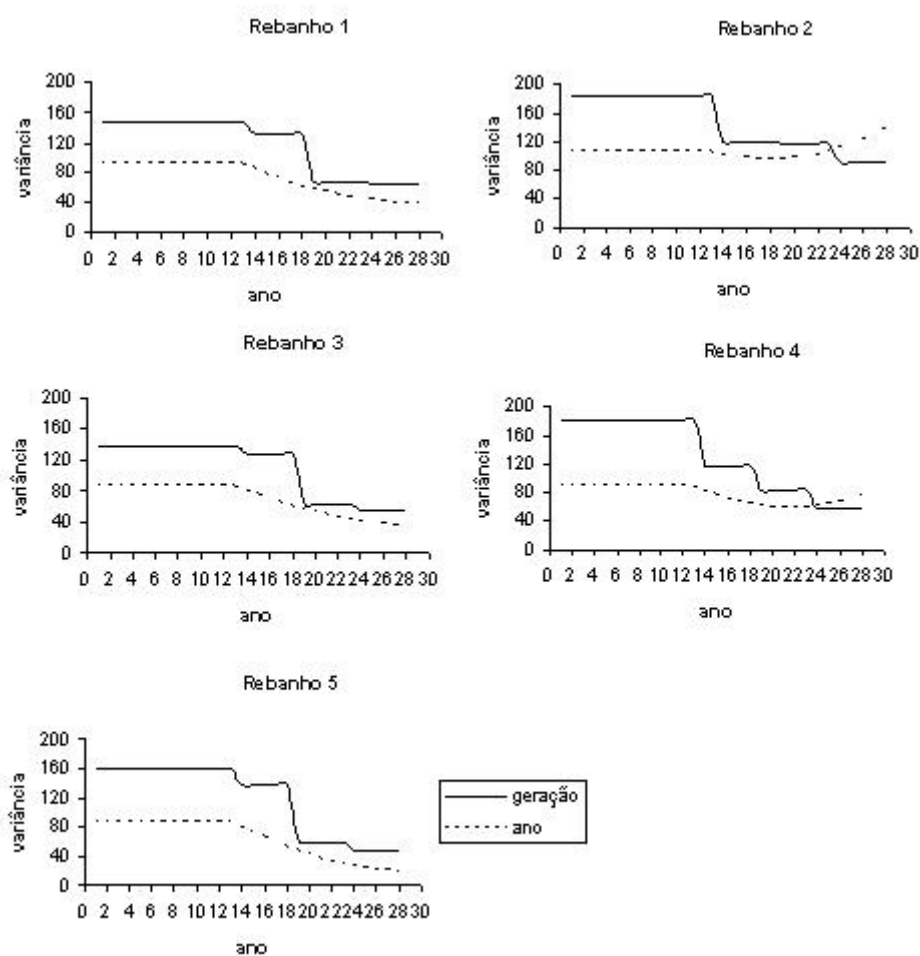


Figura 2 – Comportamento das variâncias genéticas por geração de seleção e por ano de nascimento, para P550 dias, para os rebanhos 1 ao 5.

Figure 2 – Behavior of the genetic variances for selection's generation and year of birth, for P550, for herds 1 to 5.

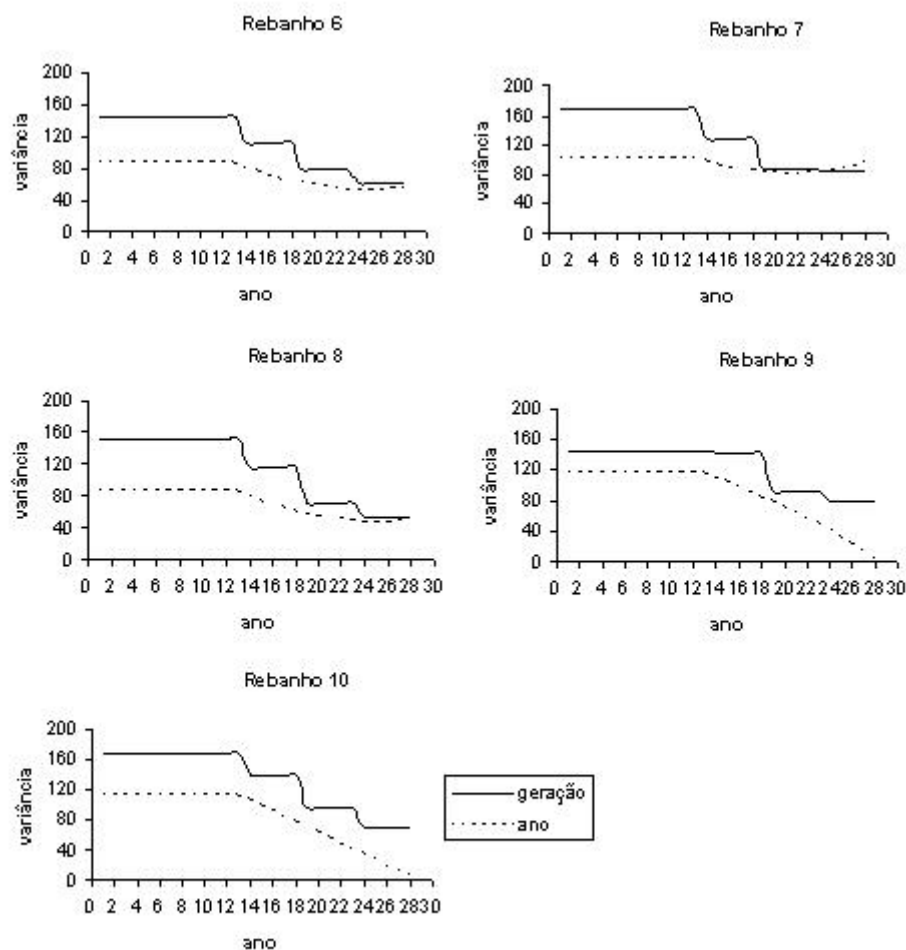


Figura 3 – Comportamento das variâncias genéticas por geração de seleção e por ano de nascimento, para P550 dias, para os rebanhos 6 ao 10.

Figure 3 – Behavior of genetic variances for selection's generation and year of birth, for P550, for herds 6 to 10.

Com o objetivo de descrever a trajetória dos componentes de variância ao longo do processo de seleção em gado de corte, Martins et al. (2002) obtiveram sucesso no uso de um modelo de regressão polinomial.

Rekaya et al. (2003) apresentaram o desenvolvimento metodológico de um modelo estrutural (ME), para estimação das (co)variâncias genéticas para produção de leite em bovinos da raça Holandesa em cinco regiões dos Estados Unidos. Neste estudo, o ME mostrou-se vantajoso nas situações em que existia número limitado de informações para estimação das (co)variâncias genéticas, principalmente, em razão dos fracos laços genéticos entre as regiões. Outra vantagem que encontraram para essa modelagem foi a possibilidade

de informações externas serem incorporadas como condições climáticas e como similaridade genética e de manejo entre as regiões, para exploração de padrões nas (co)variâncias genéticas. As estimativas de (co)variâncias genéticas e de herdabilidade, obtidas pelo ME foram muito mais precisas que as estimadas pelo modelo multicarácter, pelo primeiro modelo conter menor quantidade de parâmetros.

Falcão et al. (2006), utilizando um modelo multicarácter para a estimação de componentes de (co)variância para produção de leite em vacas da raça Holandesa em diferentes estados, usando métodos bayesianos, encontraram baixa precisão das estimativas de (co)variâncias genéticas e relacionaram isso com a escassez de informação genética disponível e grande quantidade de parâmetros a serem estimados.

Os altos valores de R^2 , aqui obtidos, apontam que a regressão foi eficientemente ajustada, e indica que os componentes de variância em cada ano e de covariância genética entre os anos podem ser estimados de forma precisa por meio da regressão múltipla, para serem usados em situações em que a heterogeneidade de (co)variância existe e deve ser contemplada.

Conclusão

Os resultados indicam que o modelo de regressão múltipla dos componentes de (co)variância em função das gerações sob seleção pode ser implementado em procedimentos bayesianos para estimação de componentes de (co)variância quando se admite heterocedasticidade ao longo do tempo, por efeito da seleção.

Literatura Citada

BOLDMAN, K.G.; KRIESE, L.A.; VAN VLECK, L.D. et al. A manual fo use of MTDFREML. A set of programs to obtain estimates of variances and covariances[Draft]. USDA-ARS, Lincoln, NE, 1995.

- COWLES, K.; BEST, N.; VINES, K. **Convergence diagnosis and output analysis**. Cambridge: MRC Biostatistics Unit, UK. Version 0.40. 1995.
- ENGLER, E. O. **Heterogeneidade de variâncias entre gerações para peso aos 365 e 550 dias de idade para animais da raça Canchim**. 2002. 51p. Dissertação (Mestrado em Zootecnia) – Universidade Estadual de Maringá, 2002.
- FALCÃO, A. J. S.; MARTINS, E. N.; COSTA, C. N. et al. Heterocedasticidade entre estados para produção de leite em vacas da raça Holandesa, usando métodos bayesianos via amostrador de Gibbs. **Revista Brasileira de Zootecnia**, v.35, n.2, p.405-414, 2006.
- FERRAZ FILHO, P.B.; RAMOS, A.A.; SILVA, L.O.C. et al. Tendência genética dos efeitos direto e materno sobre os pesos à desmama e pós-desmama de bovinos da raça Tabapuã no Brasil. **Revista Brasileira de Zootecnia**, v.31, n.2, p.635-640, 2002.
- 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 4**. Oxford: University Press, 1992. p.625-631.
- GIANOLA, D.; FOULLEY, J.L.; FERNANDO, R.L. et al. Estimation of Heterogeneous Variances Using Empirical Bayes Methods: Theoretical Considerations. **Journal of Dairy Science**, v.75, p.2805-2823, 1992.
- GOMEZ-RAYA, L.; BURNSIDE, E. B. The effect of repeated cycles of selection on genetic variance, heritability, and response. **Theoretical Applied Genetics**, v. 79, p. 568-574, 1990.
- HEIDELBERGER, P.; WELCH, P.D. Simulation run length control in the presence of an initial transient. **Operations Research**, Landing, v.31, n.6, p.1109-1144, 1983.
- HENDERSON, C. R. **Applications of linear models in animal breeding**. University of Guelph, Ontario, Canadá, 1984.
- LÔBO, R.B.; OLIVEIRA, H.N.; BEZERRA, L.F. et al. Estimativa de componentes de (co)variância herdabilidade para o peso aos 120 dias de idade na raça Nelore usando estatística bayesiana. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 34, 1997, Juiz de Fora – MG. **Anais...** Juiz de Fora: SBZ, 1997. p. 186 – 189.
- MARTINS, E.N. Avaliação Genética e Heterogeneidade de Variância. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 39, 2002, Recife-PE. **Anais...** Recife: Sociedade Brasileira de Zootecnia, [2002]. CD-ROM. Melhoramento genético animal.
- MARTINS, E.N. ; ENGLER, E.O.; SAKAGUTI, E.S. et al. In: PROCEEDINGS OF THE 7TH WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 7, 2002, Montpellier. **Anais...** Montpellier, v.32, p.303-305, 2002.
- MARTINS, E. N. Estratégias de manutenção de pequenas populações sob seleção. In: SIMPÓSIO DE PRODUÇÃO DE PERDIZES (*Rhyncatus rufescens*) EM CATIVEIRO, 1, 2003. Jaboticabal. **Anais...** Jaboticabal: FCAV-UNESP, [2003] 7 p. CD-ROM. Palestras.

- NUÑEZ-DOMINGUEZ, R.; VAN VLECK, L. D.; CUNDIFF, L. V. 1995. Prediction of genetic values of sires for growth traits of crossbred cattle using a multivariate animal model with heterogeneous variances. **J. Anim. Sci.** 73(10): 2940-2950.
- OLIVEIRA, C.A.L. **Efeito direto e materno em características de crescimento em rebanhos de bovinos de corte sob seleção.** 2006. 85p. Tese (Doutorado em Zootecnia) – Universidade Estadual de Maringá, 2006.
- PEREIRA, V. M. C. **Estudo genético de critérios de seleção ligados à eficiência reprodutiva e ao crescimento de machos e fêmeas da raça Canchim.** 2003. 70p. Dissertação (Mestrado em Zootecnia) – Universidade Estadual Paulista. 2003.
- QUINTON, M.; SMITH C. Comparison of evaluation-selection systems for maximizing genetic response at same level of inbreeding. **Journal of Animal Science**, v.73, p.2208-2212,1995.
- REKAYA, R.; WEIGEL, K.A.; GIANOLA, D. Application of a structural model for genetic covariances in international dairy sire evaluations. **Journal of Dairy Science**, v. 84, p.1525-1530, 2001.
- REKAYA, R.; WEIGEL, K.A.; GIANOLA, D. Bayesian estimation of parameters of a structural model for genetic covariances between milk yield in five regions of the United States. **Journal of Dairy Science**, v. 86, p.1837-1844, 2003.
- SIMONELLI, S.M. **Heterogeneidade de variâncias e interação genótipo x ambiente no desempenho de animais nelore em diferentes regiões do estado do Mato Grosso do Sul.** 2004. 118p. Tese (Doutorado em Zootecnia) – Universidade Estadual de Maringá, PR, 2004.
- SORENSEN, D. **Gibbs sampling in quantitative genetics.** Intern report, n.82, Danish Institute of Animal, Department of Breeding and Genetics, Denmark, 1996.
- VAN MELIS, M. H.O.; ELER, J. P.O.; SILVA, J. A. V. et al. Estimacão de parâmetros genéticos em bovinos de corte utilizando os métodos de máxima verossimilhança restrita e R. **Revista Brasileira de Zootecnia**, v.32, n.6, p.1624-1632, 2003 (suplemento).
- VAN TASSEL, C.P.; VAN VLECK, L.D. **A manual for use of MTGSAM. A set of FORTRAN programs to apply Gibbs sampling to animal models for variance component estimation (DRAFT).** Lincoln: Department of Agriculture Research Service, 1995. 86p.
- WINKELMAN, A.; SCHAEFFER, L. R. 1988. Effect of heterogeneity of variance on dairy sire evaluation. **J. Dairy Sci.** 71(11):3033-3039.

IV – Métodos para estimação de tendência genética

RESUMO – Foram estudados métodos alternativos para a estimação da tendência genética para peso aos 550 dias em 10 rebanhos de bovinos de corte simulados, submetidos há 20 anos de seleção, considerando a existência de parentesco entre os animais e contemplando a heterogeneidade de (co)variância pela inclusão da estrutura de (co)variância genética entre os anos de nascimento. Para a construção da estrutura de (co)variância foi utilizado um modelo de regressão múltipla a partir dos componentes estimados para cada geração de seleção, obtidos por análise multicarácter implementada para inferência bayesiana. As tendências genéticas foram estimadas para 8 anos de seleção, com duas condições (C1 e C2) aplicadas à avaliação genética; realizada na C1 a avaliação, em todos os animais até o 20º ano de seleção, e na C2, apenas, os animais que participaram do conjunto de dados utilizado para fazer a estimação da tendência genética. Esta foi obtida por meio de equações de regressão ajustadas, alternativamente, pelos métodos de quadrados mínimos ordinários (OLS), quadrados mínimos ponderados (WLS) e quadrados mínimos generalizados (GLS), sob duas formas: uma considerando homogeneidade de variância, e a outra, heterogeneidade de variância ao longo dos anos de seleção. Foi verificado comportamento quadrático dos valores genéticos em função do ano de nascimento dos animais, para os 10 rebanhos estudados na condição 1 e para os rebanhos 1, 4, 8 e 9 na condição 2. Para a C1, os métodos WLS e GLS, apesar de permitirem maior inclusão de informações no modelo, não foram capazes de detectar as mudanças genéticas com maior precisão, sendo o OLS o método que permitiu maior aproximação às médias anuais dos valores genéticos reais e preditos. Para a C2, o método GLS, considerando heterogeneidade de variância ao longo dos anos de seleção, foi o que mais se aproximou das mudanças genéticas reais.

Palavras-chave: regressão múltipla, heterogeneidade de variância, inferência Bayesiana

IV – Methods for estimation of genetic trends

ABSTRACT – Alternative methods for the estimation of genetic trends for weight at 550 days (P550) in 10 beef herds, submitted to 20 years of selection were studied. The relationship between animals and the heteroscedasticity of genetic (co)variance structure between the year of birth of the animals were contemplated. For the construction of the structure of (co)variance a model of multiple regression from the components estimated for each generation of selection was used. The components were obtained from multi-character analysis implemented for Bayesian inference. The genetic trends were estimated for 8 years of selection, with two conditions (C1 and C2) applied to the genetic evaluation. In C1 the evaluation considered all the animals until 20th year of selection and in C2 only the animals that were included on the data set used to make the estimation of the genetic trend were considered. The genetic trend was obtained by the adjustment of regression equations, alternatively, for the methods of Ordinary Least Squares (OLS), Weighted Least Squares (WLS) and Generalized Least Squares (GLS), the latter under two forms, one considering homogeneity of variance and another considering heterogeneity of variance throughout the years of selection. A quadratic behavior of the genetic values as a function of the year of birth of the animals was verified for the 10 herds studied in the first condition and for herds 1, 4, 8 and 9 in the second condition. For C1 WLS and GLS methods, although allowing greater inclusion of information in the model, were not able to detect the genetic changes with greater precision, being the OLS the method that allowed a greater approach to the annual averages of the real and predicted genetic values. For C2, the GLS method considering heterogeneity of variance throughout the years of birth of the animals was closer to the real genetic changes.

Key words: multiple regression, heterogeneity of variance, Bayesian inference

Introdução

Programas de melhoramento genético são fundamentais para alcançar melhorias na genética dos rebanhos. Porém, depois do programa estar implantado, é necessário que seja feita, periodicamente, a verificação e acompanhamento de sua eficiência (Mello, 1999). Uma das maneiras de realizar um acompanhamento é por meio do conhecimento de parâmetros genéticos e de estimativas de mudança genética, imprescindíveis para o estabelecimento de diretrizes que guiem os programas de melhoramento, avaliando o progresso genético ao longo do tempo para que os resultados sirvam de elementos orientadores para ações futuras (Euclides Filho et al., 1997).

A tendência genética é uma medida que permite avaliar a mudança ocasionada por um processo de seleção, tornando possível a quantificação da porção genética responsável pelas mudanças acumuladas ao longo dos anos, em determinado rebanho. A estimativa da tendência genética é, até então, a melhor maneira de se observar o progresso genético, visto que a melhora no desempenho não significa obrigatoriamente melhoria genética (Zollinger & Nielsen, 1984; Euclides Filho et al., 1997).

Vários autores utilizam essa metodologia para estudar o comportamento do valor genético ao longo dos anos, dentre eles: Euclides Filho et al. (1997) em rebanhos Guzerá, Ferraz Filho et al. (1997) em animais da raça Nelore Mocha, Fernandes et al. (2002) na raça Charolês, Silva et al. (1997) e Holanda et al. (2004) em rebanhos Nelore.

A tendência genética é obtida pela regressão dos valores genéticos ao longo do tempo e os coeficientes da regressão são estimados por meio do método de quadrados mínimos.

De acordo com Martins et al. (1997) e Oliveira (1998), o método dos quadrados mínimos é dito ordinário (Ordinary Least Square - OLS) quando se considera que os

resíduos não são correlacionados e há homocedasticidade. Em casos em que existe heterocedasticidade, porém com resíduos não-correlacionados, o método é dito de quadrados mínimos ponderados (Weighted Least Square - WLS). Se existe correlação entre os resíduos, havendo ou não homogeneidade de variância, pode-se estimar os parâmetros pelo método dos quadrados mínimos generalizados (Generalized Least Square - GLS).

Alguns autores (Ferraz Filho et al., 1997, 2002; Fernandes et al., 2002; Holanda et al., 2004), ao utilizarem o método de quadrados mínimos ordinários e fazendo o ajuste linear dos valores genéticos em função do ano de nascimento do animal, encontraram taxas de mudança genética anual aquém do potencial sugerido por Smith (1985), que é de 1 a 3% da média da população. Todavia, o método de quadrados mínimos ordinários apresenta propriedades ótimas somente quando os resíduos são não-correlacionados, a distribuição residual é normal e as variâncias são homogêneas, o que nem sempre é a realidade, em função do parentesco entre os animais e do impacto da seleção sobre os parâmetros genéticos.

Em vista disso, o objetivo deste trabalho foi avaliar métodos para a estimação de tendências genéticas, comparando métodos usuais e métodos que contemplem a estrutura de (co)variância entre os valores genéticos ao longo das gerações e o parentesco existente entre os animais.

Material e Métodos

O conjunto de dados utilizado, formado por rebanhos de bovinos de corte submetidos a 20 anos de seleção, foi simulado em linguagem de programação Fortran, compilador F90.

Nesses rebanhos, o plantel de reprodução foi formado de 1.500 vacas e 38 touros, mantendo-se na monta natural uma relação touro-vaca de 1:40. A partir do segundo ano, foi implementado inseminação artificial em 50% das vacas, e utilizados 10 touros para inseminação artificial e 19 touros para monta natural. A taxa de natalidade estabelecida foi de 90% e sobrevivência de 95% até o início da reprodução.

A taxa de descarte de vacas era variável e determinada pelo número de vacas vazias ao final da estação monta; apenas as vacas primíparas poderiam permanecer no plantel de reprodução se estivessem vazias. Os animais jovens poderiam ser incluídos no plantel de reprodução, a partir de 22 meses de idade. Foram simulados dados de peso ao desmame (PD) aos 550 dias (P550), assim como os valores genéticos para efeito direto de peso ao desmame e peso aos 550 dias e efeito materno de peso ao desmame.

A cada ano procedeu-se a avaliação genética com o objetivo de orientar a seleção, utilizando-se as equações de modelos mistos (Henderson, 1984) para predição dos valores genéticos, por meio de análise bicaracter, utilizando-se o sistema computacional MTDFREML (*Multiple Trait Derivative Free Restricted Maximum Likelihood*) (Boldman et al. 1995). Os efeitos identificáveis de ambiente foram: sexo (macho ou fêmea), época de nascimento (início, meio ou final da estação de nascimento), ano de nascimento e idade (em meses) da mãe ao parto. Os valores dos componentes de (co)variância genética inicialmente utilizados foram os seguintes,

$$G_0 = \begin{bmatrix} \sigma_{a_D}^2 & \sigma_{a_{Dm}} & \sigma_{a_{DS}} \\ \sigma_{a_{Dm}} & \sigma_m^2 & \sigma_{a_{Sm}} \\ \sigma_{a_{DS}} & \sigma_{a_{Sm}} & \sigma_{a_S}^2 \end{bmatrix} = \begin{bmatrix} 165 & -55 & 135 \\ -55 & 95 & 0 \\ 135 & 0 & 215 \end{bmatrix}, \text{ em que,}$$

$\sigma_{a_i}^2$ é a variância genética aditiva para característica i , $i=PD(D)$ e $P550(S)$;

σ_m^2 é a variância genética aditiva materna da característica PD;

$\sigma_{a_{DS}}$ é a covariância genética entre os efeitos diretos de PD e P550;

$\sigma_{a,m}$ é a covariância genética entre os efeitos diretos das características i com os efeitos maternos de PD, $i=PD(D)$, P550(S);

As variâncias residuais adotadas foram 365 e 450 kg² para PD e P550, respectivamente, e o valor da covariância residual foi de 150 kg².

Para classificação dos candidatos à seleção utilizou-se um índice empírico de seleção com pesos de 0,3; 0,3; 0,4, respectivamente para os valores genéticos direto e materno para PD e direto para P550.

Os acasalamentos dos animais selecionados foram realizados ao acaso, com restrição para controlar o incremento dos níveis de endogamia, por meio impedimento do acasalamento entre pais e filhos e entre irmãos (meio-irmãos ou irmãos completos).

Após o estabelecimento das condições de simulação e seleção, foram realizadas 10 repetições, totalizando 10 rebanhos com média de 31198 animais, conforme Oliveira (2006).

Apenas a característica P550 foi estudada, e a edição dos dados foi feita por meio de programas e desenvolvida em linguagem Fortran.

Para a obtenção das (co)variâncias heterogêneas, para a característica peso aos 550 dias, primeiramente, foi realizada a estimação dos componentes de variância de P550 de cada geração de seleção (5 anos) e covariância de P550 entre as gerações, por meio do software MTGSAM (*Multiple Trait Gibbs Sampling in Animal Model*), em análise multicarácter em que o peso aos 550 dias foi tratado como 4 características distintas de acordo com a geração em que o animal nasceu, sendo P550-g1 o peso aos 550 dias para os animais nascidos do ano 1 ao ano 5, P550-g2 para os nascidos entre 6 e 10, P550-g3 para 11 ao 15 e P550-g4 para os nascidos entre os anos 16 e 20. Foi considerado também que um animal só possui observação para a característica referente

à classe a que ele pertence, ao apresentar dados perdidos nas características referentes às demais classes. Sexo, estação de nascimento e ano de nascimento foram considerados efeitos fixos, e idade da mãe ao parto foi considerada como covariável.

Embora o MTGSAM seja um programa flexível quanto aos modelos de análise que podem ser utilizados, as distribuições *a priori* dos efeitos considerados no modelo são definidas pelo programa, e necessárias para definir apenas alguns parâmetros da distribuição dos componentes de variância (Lôbo et al., 1997).

Assim, foi assumido que para os efeitos de ambiente identificáveis não existe nenhum conhecimento inicial, tendo uma distribuição inicial uniforme e que a distribuição dos efeitos genéticos e dos resíduos é normal multivariada. Em se tratando dos valores genéticos, foi considerada a estrutura de (co)variância conhecida pela matriz de parentesco. Os *priors* para variância genética aditiva foram 165, 110, 75 e 70 kg² para P550-g1, P550-g2, P550-g3 e P550-g4, respectivamente; para a variância residual foi 450 kg² para as 4 características. As definições para os elementos do modelo animal e da distribuição conjunta de Y , a e e são, respectivamente:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} X_1 & 0 & 0 & 0 \\ 0 & X_2 & 0 & 0 \\ 0 & 0 & X_3 & 0 \\ 0 & 0 & 0 & X_4 \end{bmatrix} \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} + \begin{bmatrix} Z_1 & 0 & 0 & 0 \\ 0 & Z_2 & 0 & 0 \\ 0 & 0 & Z_3 & 0 \\ 0 & 0 & 0 & Z_4 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \end{bmatrix};$$

em que:

y_1, y_2, y_3, y_4 são os vetores de observações, para peso aos 550 dias de idade, referentes às gerações 1, 2, 3 e 4 respectivamente;

X_1, X_2, X_3, X_4 são as matrizes de incidência dos efeitos identificáveis de ambiente, para peso aos 550 dias de idade, referentes às gerações 1, 2, 3 e 4, respectivamente;

$\beta_1, \beta_2, \beta_3, \beta_4$ são os vetores dos efeitos identificáveis de ambiente, para peso aos 550 dias de idade, referentes às gerações 1, 2, 3 e 4, respectivamente;

Z_1, Z_2, Z_3, Z_4 são as matrizes de incidência dos efeitos aleatórios, para peso aos 550 dias de idade, referentes às gerações 1, 2, 3 e 4, respectivamente;

a_1, a_2, a_3, a_4 são os vetores dos efeitos genéticos diretos, para peso aos 550 dias de idade, referentes às gerações 1, 2, 3 e 4, respectivamente;

e_1, e_2, e_3, e_4 são os vetores dos erros aleatórios, para peso aos 550 dias de idade, referentes às gerações 1, 2, 3 e 4, respectivamente;

com distribuição conjunta:

$$\begin{bmatrix} y \\ a \\ e \end{bmatrix} \sim NMV \left\{ \begin{bmatrix} X\beta \\ 0 \\ 0 \end{bmatrix}; \begin{bmatrix} ZGZ'+R & ZG & R \\ GZ' & G & 0 \\ R & 0 & R \end{bmatrix} \right\}, \text{ em que:}$$

$$G = G_0 \otimes A$$

\otimes é o produto de Kronecker;

G_0 é a matriz de (co)variância genética das i gerações, para peso aos 550 dias de idade, dada a seguir:

$$G_0 = \begin{bmatrix} \sigma_{a_1}^2 & \sigma_{a_1a_2} & \sigma_{a_1a_3} & \sigma_{a_1a_4} \\ \sigma_{a_1a_2} & \sigma_{a_2}^2 & \sigma_{a_2a_3} & \sigma_{a_2a_4} \\ \sigma_{a_3a_1} & \sigma_{a_3a_2} & \sigma_{a_3}^2 & \sigma_{a_3a_4} \\ \sigma_{a_4a_1} & \sigma_{a_4a_2} & \sigma_{a_4a_3} & \sigma_{a_4}^2 \end{bmatrix}, \text{ sendo:}$$

$\sigma_{a_i}^2$ é a variância genética aditiva da geração i ; $i = 1, 2, 3$ e 4

$\sigma_{a_i a_j}$ é a (co)variância genética entre as gerações i e j ;

$i = 1, 2, 3$ e 4 e $j = 1, 2, 3$ e 4

$$R = R_0 \otimes I$$

em que:

I é a matriz identidade de ordem igual ao número de animais;

R_0 é a matriz de variância residual, das i gerações, para peso aos 550 dias de idade, dada a seguir:

$$R_0 = \begin{bmatrix} \sigma_{e_1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{e_2}^2 & \dots & 0 \\ \vdots & \vdots & \sigma_{e_3}^2 & \vdots \\ 0 & 0 & \dots & \sigma_{e_4}^2 \end{bmatrix}, \text{ em que:}$$

$\sigma_{e_i}^2$ é a variância residual da geração i ; $i = 1, 2, 3$ e 4 .

Para os efeitos identificáveis de ambiente foi assumida distribuição “uniforme”.

Para os componentes de (co)variância genética foi assumido que G têm distribuição de Wishart Invertida (IW), na forma

$$G \sim IW(G_0, \nu_g),$$

tal que

$$f(G/G_0, \nu_g) \propto |G|^{-\frac{1}{2}(\nu_g + m_g + 1)} \cdot e^{-\frac{1}{2}tr(G_0^{-1}G^{-1})}$$

em que:

ν_g é o grau de liberdade da distribuição que equivale ao grau de crença que se tem acerca do parâmetro,

m_g é a ordem da matriz G ,

G_0 é a matriz de parâmetro escala da distribuição de G .

Para os componentes de variância residual foi assumido que cada elemento $\sigma_{e_i}^2$ de R_0 , tem distribuição qui-quadrado invertida, na forma

$$\sigma_{e_i}^2 \sim IX^2 (S_0, \nu_r)$$

tal que:

$$f(\sigma_{e_i}^2/s_0, \nu_r) \propto (\sigma_{e_i}^2)^{-\frac{1}{2}(\nu_r + 2)} \cdot e^{-\frac{1}{2}(s_0/\sigma_{e_i}^2)}$$

em que:

v_{ri} é o grau de liberdade da distribuição que equivale ao grau de crença que se tem acerca dos parâmetro,

s_{0i} é o parâmetro escala da distribuição de $\sigma_{e_i}^2$.

São calculados como segue:

$$G_0 = \frac{G^*}{v_g^*} = \frac{G^*}{(v_g - m_g - 1)}$$

em que G^* é a matriz de *priors* dos componentes de variância contidos em G .

$$s_{0i} = \frac{\sigma_{e_i}^{2*}}{v_{r_i}^*} = \frac{\sigma_{e_i}^{2*}}{(v_{r_i} - 2)},$$

em que $\sigma_{e_i}^{2*}$ é o *prior* do componente de variância residual $\sigma_{e_i}^2$.

A função densidade de probabilidade conjunta posterior dos parâmetros, dados os hiperparâmetros, foi obtida por meio do produto das distribuições *a priori* pela função de verossimilhança, conforme descrito por Van Tassel & Van Vleck (1995). Para implementar a amostragem de Gibbs, foram obtidas as funções densidade de probabilidade condicionais completas para cada parâmetro, por meio da derivação da função de densidade posterior conjunta, considerando como constantes os valores dos outros parâmetros contidos no modelo, que gera a distribuição do parâmetro de interesse:

$$\begin{aligned} & f(\beta, u, G, R \mid y, v_g, G_0, v_{e_1}, s_{0_1}, \dots, v_{e_4}, s_{0_4}) \\ & \propto |R|^{-1/2} \cdot \exp\left\{-\frac{1}{2}(y - X\beta - Zu)' R^{-1}(y - X\beta - Zu)\right\} \\ & \cdot |G|^{-\frac{1}{2}(n+v_g+m_g+1)} \cdot \exp\left\{-\frac{1}{2}[a'(G^{-1} \otimes A^{-1})a + tr(v_g^* G_0 G^{-1})]\right\} \\ & \prod_{i=1}^7 \left[(\sigma_{e_i}^2)^{-1/2}(q_i+v_{e_i}+2) \cdot \exp\left\{-\frac{v_{r_i}^*}{2} s_{0_i} \sigma_{e_i}^{-2}\right\} \right]. \end{aligned}$$

As estimativas das médias *posteriores* para os componentes de (co)variância foram baseadas no valor esperado de suas distribuições condicionais completas.

Para cada uma das repetições (rebanhos) foi gerada uma cadeia de Gibbs de 1.000.000 de ciclos com amostras, retiradas a cada 100 ciclos, após a eliminação de 50.000 ciclos iniciais, gerando cadeias de 9.500 amostras dos componentes de (co)variância. Foi diagnosticada a convergência das cadeias de amostragem de Gibbs ao utilizar os métodos de Geweke (1992) e de Heidelberger & Welch (1983), disponíveis na biblioteca CODA (Convergence Diagnosis and Output Analysis, versão 0.4), desenvolvida por Cowles et al. (1995), e implementada no software R (2004).

Então, a fim de obter as (co)variâncias genéticas para P550 entre e para os anos em que realizou a seleção, foi utilizado um modelo de regressão múltipla em que:

$$Y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1^2 + b_4 x_2^2 + b_5 x_1 x_2, \text{ em que:}$$

Y é o componente de (co)variância que se pretende estimar;

b_0 é a constante geral;

b_1, b_2, b_3, b_4 e b_5 são os coeficientes da regressão;

x_1 e x_2 são os anos de nascimento dos animais.

A estimação dos coeficientes da regressão múltipla foi realizada pelo método dos quadrados mínimos generalizados (GLS), conforme:

$$\beta = (X'V^{-1}X)^{-1}X'V^{-1}y, \text{ em que:}$$

β é o vetor dos coeficientes da regressão;

X é a matriz de incidência dos efeitos fixos de ano;

V é a matriz de (co)variância entre os componentes estimados pelo software MTGSAM;

y é o vetor de médias posteriores dos componentes estimados.

Para cada rebanho foi ajustada uma equação que permitiu a obtenção da matriz de (co)variância genética de P550 entre os anos de nascimento dos animais. Os valores dos

componentes de (co)variância genética para P550 do último ano da primeira geração de seleção (ano 5) foram atribuídos a todos os anos dessa primeira geração.

De posse das estimativas de variância e covariância genéticas, para peso aos 550 dias de acordo com o ano de nascimento do animal e dos valores genéticos preditos pela avaliação genética dos animais, foi possível testar a eficiência de quatro diferentes métodos de estimação do comportamento dos valores genéticos ao longo dos anos em que os rebanhos foram submetidos à seleção. Para isso, foram desenvolvidos programas em linguagem Fortran, e usado compilador F90.

Por causa de limitações computacionais em função do grande número de animais nos rebanhos (em média 31.198, distribuídos em 28 anos – 1996 a 2023), somente foi possível a estimação da tendência genética em um intervalo de 8 anos de seleção, que engloba em média 10.778 animais por rebanho. A edição dos dados foi realizada por programas desenvolvidos em linguagem Fortran, e retirados os animais nascidos antes do ano 9 (2004) e após o ano 16 (2011). Os animais foram renumerados, assim como os anos de nascimento.

Para a estimação dos coeficientes de regressão das tendências genéticas, nos quatro métodos, foi utilizado o seguinte modelo:

$$\hat{a} = X\beta + e, \text{ em que:}$$

\hat{a} é o vetor dos valores genéticos preditos pela avaliação genética;

X é a matriz de incidência das constantes contidas no vetor β ;

β é o vetor dos coeficientes da regressão, sendo $\beta = [b_0 \ b_1 \ b_2]$;

e é o vetor de erros aleatórios.

A solução para o vetor β foi obtida conforme a expressão

$$\beta = (X'G^{-1}X)^{-1}X'G^{-1}\hat{a}, \text{ sob quatro diferentes métodos:}$$

No primeiro método, para a obtenção de β , o vetor dos valores genéticos preditos (\hat{a}) foi considerado como sendo independentemente distribuído, ou seja, sem correlação entre os resíduos, com variância homogênea ao longo das gerações de seleção, sendo:

$$G = I\sigma_a^2$$

Como a variância genética aditiva, σ_a^2 , neste caso é homogênea, ou seja, é um escalar que pondera os valores genéticos tanto na primeira $((X'G^{-1}X)^{-1})$ como na segunda $(X'G^{-1}\hat{a})$ parte da equação, ele foi eliminado, ficando:

$\beta = (X'X)^{-1}X'\hat{a}$, que é a forma com que a tendência genética é, normalmente, estimada. Essa metodologia recebeu o nome de TG1 e foi implementada em linguagem Fortran, assim como os 3 métodos abaixo descritos.

Foi assumido no segundo método, que o vetor \hat{a} é independentemente distribuído, mas existe heterogeneidade de variância decorrente do ano em que o animal nasceu, sendo a matriz $G = I\sigma_{a_i}^2$, em que $\sigma_{a_i}^2$ é a variância genética aditiva no ano i . Constituindo o método TG2.

Para o terceiro método, denominado TG3, foi admitido que o vetor \hat{a} apresenta uma estrutura de correlação, considerando então a presença de efeitos aleatórios para a obtenção dos estimadores da tendência genética, por meio da inclusão da matriz de parentesco no modelo. A heterogeneidade genética não foi contemplada neste método, e G representado por $A\sigma_a^2$, sendo a variância genética aditiva, σ_a^2 , um escalar, a estimação dos coeficientes da regressão foi feita por meio de:

$$\beta = (X'A^{-1}X)^{-1}X'A^{-1}\hat{a}, \text{ em que } A \text{ é a matriz de parentesco entre os animais.}$$

No quarto método, a estimação dos coeficientes da regressão do valor genético em função do ano de nascimento do animal, foi realizada levando em consideração a presença de efeitos aleatórios no modelo, representados pela matriz de parentesco entre os animais, que estabelecem estrutura de correlação entre eles. Foi considerada,

também, a existência de variâncias heterogêneas, ocasionadas pelo processo de seleção ao longo dos anos. A matriz G , nesse caso, foi representada por $A\sigma_{a,a_j}$, em que A é a matriz de parentesco entre os animais e σ_{a,a_j} é a (co)variância genética entre os anos i e j . A estrutura de (co)variância utilizada foi editada a partir da estrutura original, restando apenas os componentes representantes dos anos 9 ao 16. Essa metodologia recebeu o nome de TG4.

Assim, foram ajustadas equações de regressão na forma

$$\hat{a}_{ij} = b_0 + b_1x_j + b_2x_j^2,$$

em que:

\hat{a}_{ij} é o valor genético predito, por cada um dos quatro métodos, dos efeitos diretos de P550 do animal i nascido no ano j ;

b_0 é constante geral;

b_1 é o coeficiente linear de regressão de \hat{a}_{ij} em função do ano de nascimento;

b_2 é o coeficiente quadrático de regressão de \hat{a}_{ij} em função do ano de nascimento;

x_j é o ano de nascimento j .

Com a intenção de verificar o impacto da estrutura de informação de família na avaliação genética e na estimação da tendência genética, foram testadas duas condições, C1 e C2, para cada um dos quatro programas. Na condição C1, foi utilizada a avaliação genética com todos os animais nascidos dentro dos 28 anos, ou seja, acumulados nas 4 gerações de seleção, conforme descrito anteriormente. Na condição C2, a avaliação genética foi realizada com base em animais que fariam parte da estimação da tendência genética, ou seja, os pertencentes a duas gerações de seleção.

Para as duas condições, os valores genéticos para peso aos 550 dias foram preditos por meio do sistema computacional MTDFREML (*Multiple Trait Derivative Free Restricted Maximum Likelihood*) (Boldman et al., 1995). Os efeitos identificáveis

de ambiente foram: sexo (macho ou fêmea), época de nascimento (início, meio ou final da estação de nascimento), ano de nascimento e idade (em meses) da mãe ao parto. A variância genética aditiva considerada para P550 foi 215 kg² e para variância residual foi 450 kg².

A qualidade do ajustamento foi avaliada por meio do coeficiente de determinação (R^2) das curvas obtidas em cada um dos quatro métodos.

As equações ajustadas também foram testadas quanto à capacidade de prever as médias anuais dos valores genéticos reais simulados (VGS550) por meio da soma dos desvios ao quadrado entre os VGS550 e os valores genéticos preditos pelas equações ajustadas em cada um dos 4 métodos.

Todas as equações ajustadas foram testadas quanto à significância dos parâmetros por meio de análise de variância da regressão, usando teste F no nível de 5%.

Resultados e Discussão

As médias dos valores genéticos preditos por ano podem ser vistas na Tabela 1 para a condição 1 e na Tabela 2 para a condição 2. Foi observada a subestimação dos valores genéticos para C2 em relação a C1, fato que é explicado pela maior informação de parentesco utilizada na predição dos valores genéticos na condição 1. Houve também subestimação dos valores genéticos de ambas as condições em relação aos valores genéticos reais, ou seja, os que foram simulados juntamente com o conjunto de dados.

Tabela 1– Média dos valores genéticos preditos por ano, para a condição 1 (C1)
 Table 1– Means of the yearly predicted genetic values for the first condition (C1)

Rebanhos <i>Herds</i>	Ano <i>Year</i>							
	1	2	3	4	5	6	7	8
1	-1,3497	2,1244	4,0897	5,9343	8,8374	10,9601	12,0293	13,2270
2	-2,1920	2,1107	6,0088	8,1012	10,3575	11,5033	12,7114	14,8595
3	-2,3222	1,2582	4,2294	5,3963	8,8266	9,9310	12,0355	13,5234
4	-2,4437	2,4541	4,3097	6,5321	9,1826	11,1102	12,2021	14,2659
5	-2,0727	1,1348	4,1633	6,5340	8,3970	10,6698	12,3227	13,9567
6	-1,2611	1,4621	3,5474	6,3784	9,7058	10,7813	12,0808	14,5947
7	-2,3672	2,3205	5,6265	7,9134	9,7090	12,2730	14,7835	16,6092
8	-0,7716	2,9123	5,8051	7,5336	10,9565	12,1301	13,9971	15,9559
9	-1,9236	1,8993	5,7973	8,4133	10,8752	12,5901	14,4715	17,0783
10	-1,8872	0,9357	3,9943	7,3276	9,9050	11,8514	13,6481	15,5751

Tabela 2– Média dos valores genéticos preditos por ano, para a condição 2 (C2)
 Table 2 – Means of the yearly predicted genetic values for the second condition (C2)

Rebanhos <i>Herds</i>	Ano <i>Year</i>							
	1	2	3	4	5	6	7	8
1	-0,6123	-0,0958	0,1125	-0,5800	-0,6360	-0,2943	0,6878	1,0864
2	-0,1817	-0,3487	-0,3892	-0,2638	-0,6861	0,0968	-0,4046	-0,4865
3	-0,2204	-0,1210	-0,0320	-0,4724	-0,5853	-0,5370	-0,3914	-0,9560
4	0,0004	-0,2332	-0,2344	0,3273	-0,5636	0,0279	-0,2429	-2,6637
5	-0,9637	-0,2023	-0,2000	-0,3537	-1,8030	-0,5280	-0,9233	-0,2428
6	-0,2374	-0,2840	-0,2269	-0,0945	-0,3594	0,1083	0,4215	0,6115
7	-0,1343	0,0368	-0,2395	-0,5733	-0,3679	1,4067	0,4932	0,1911
8	0,0158	-0,1619	-0,2426	0,4955	-0,4133	0,0107	0,6058	1,2334
9	-0,8001	-0,2596	-0,3441	0,2122	-0,3299	0,2594	-0,5826	-0,6762
10	-1,0835	0,0056	-0,2326	-0,6793	-1,7349	-2,5027	-1,6380	-2,1499

A mudança média, em kg por ano, para os valores genéticos reais, preditos na C1 e preditos na C2, dentro do período estudado, pode ser visto na Tabela 3 para todos os rebanhos. A mudança genética média observada na C1 foi mais aproximada à mudança genética média real. Nos rebanhos 2, 3, 4 e 10 foi observada redução dos valores genéticos na condição 2, que pode ser relacionada à carência de informações de parentesco. Alguns autores como Fernandes et al. (2002) relacionam tendências genéticas negativas à falta de um programa de melhoramento genético bem estruturado,

no entanto, neste caso isso não é plausível, pois para a C1 foi utilizado igual critério de seleção e os valores para esses rebanhos apresentaram-se positivos.

Tabela 3 – Mudança genética média em kg por ano, para valor genético real, C1 e C2
 Table 3 – Mean genetic change in kg per year, for real genetic value, C1 and C2

Rebanhos <i>Herds</i>	Mudança Genética <i>Genetic change</i>		
	Real	C1	C2
1	2,83973	1,82209	0,21234
2	3,58192	2,13144	-0,03810
3	2,37936	1,98070	-0,09195
4	3,34774	2,08871	-0,33301
5	2,76368	2,00368	0,09012
6	3,10176	1,98198	0,10612
7	3,69688	2,37205	0,04068
8	2,77117	2,09094	0,15221
9	3,62671	2,37523	0,01548
10	3,74075	2,18279	-0,13330

Ferraz Filho et al. (1997), ao avaliar as mudanças genéticas em rebanhos Nelore Mocho, obtiveram tendências positivas para peso à desmama com ganhos de 1,5 kg por ano. Silva et al. (1997), ao estudar a tendência genética em animais da raça Nelore, encontraram incremento anual de 0,259 kg para peso pré-desmama. Fernandes et al. (2002) encontraram tendências genéticas de -0,0105 kg anuais para peso ao nascimento e de 0,041 kg anuais para ganho de peso do nascimento à desmama, em animais da raça Charolês.

De acordo com Kennedy et al. (1988), os valores genéticos dependem da estimativa de herdabilidade na população base, assim, a resposta será subestimada se a herdabilidade for menor que o verdadeiro valor e será superestimada se a herdabilidade for maior. O entendimento e a definição correta da população base são critérios de uma análise considerando modelo animal. Com a matriz de parentesco, animais sem pais identificados são tratados como animais da população base com a suposição de que

foram amostrados de uma única população com valor genético médio nulo e de mesma variância genética aditiva.

A partir da análise de regressão dos valores genéticos preditos, em função do ano de nascimento dos animais, ficou evidente para a condição 1 (C1) que a tendência genética apresentou comportamento quadrático, com incremento dos valores genéticos e redução dos ganhos genéticos ao longo dos anos de seleção, o que já era esperado, pois em populações fechadas sob seleção, a variância genética diminui ao longo das gerações de seleção, fazendo com que o ganho genético seja cada vez menor. Na condição 2 (C2), a tendência genética apresentou comportamento quadrático para os rebanhos 1, 4, 8 e 9; comportamento linear para os rebanhos 3, 6, 7 e 10. Para os rebanhos 2 e 5, não houve efeito do ano sobre o valor genético dos animais.

Oliveira (2006), trabalhando com bovinos de corte em rebanhos simulados, encontrou comportamento quadrático das tendências genéticas estimadas, a partir de valores genéticos reais e preditos para as características peso ao desmame e peso ao sobreano e para efeito materno de peso ao desmame.

Silva et al. (1997), em animais da raça Nelore, Ferraz Filho et al. (2002), em bovinos da raça Tabapuã, Fernandes et al. (2002), na raça Charolês, utilizaram a metodologia de tendência genética para estudar o comportamento do valor genético, no entanto, os autores consideraram apenas o efeito linear do ano sobre o valor genético, indicando haver uma única taxa de mudança genética anual para todo o período estudado.

As estimativas dos componentes de variância genética para ano de nascimento, utilizadas pelos métodos TG2 e TG4, para os anos em que foi estudado o comportamento dos valores genéticos, podem ser vistas na Tabela 4, para todos os

rebanhos. Na Tabela 5, estão os valores de covariância, utilizados pelo método TG4, para todos os rebanhos.

Tabela 4 – Componentes de variância genética aditiva estimados por ano para P550
 Table 4 – Components of additive genetic variance estimated by year, for P550

Rebanhos <i>Herds</i>	Componentes <i>Components</i>			
	$\sigma_{a_{1...5}}^2$	$\sigma_{a_6}^2$	$\sigma_{a_7}^2$	$\sigma_{a_8}^2$
1	91,33903	84,90835	78,88010	73,25426
2	106,18484	102,11155	98,93034	96,64122
3	86,60608	80,74418	75,22620	70,05212
4	89,84806	82,96075	76,91037	71,69693
5	87,86214	80,02488	72,64291	65,71625
6	85,47890	80,21008	75,40450	71,06213
7	103,02475	97,85220	93,34388	89,49979
8	86,49967	80,31046	74,64462	69,50215
9	116,05925	110,76962	105,18190	99,29606
10	113,51391	106,53682	99,53407	92,50564

Tabela 5 – Componentes de covariância genética aditiva estimados por ano para P550
 Table 5 – Components of additive genetic variance estimated by year, for P550

Rebanhos <i>Herds</i>	Componentes <i>Components</i>						
	$\sigma_{a_{1...5}a_{1...5}}$	$\sigma_{a_{1...5}a_6}$	$\sigma_{a_{1...5}a_7}$	$\sigma_{a_{1...5}a_8}$	$\sigma_{a_6a_7}$	$\sigma_{a_6a_8}$	$\sigma_{a_7a_8}$
1	91,33903	75,58481	61,60284	49,39311	69,84147	56,54684	64,50056
2	106,18484	91,41413	77,83411	65,44476	87,64512	74,36935	84,76817
3	86,60608	71,53912	58,17305	46,50788	66,35640	53,66952	61,51760
4	89,84806	75,55906	62,69807	51,26510	69,36311	57,19348	64,00409
5	87,86214	71,21255	56,29911	43,12185	64,04747	49,80622	57,33769
6	85,47890	71,67175	59,43660	48,77346	67,00462	55,37115	62,80071
7	103,02475	88,96375	76,09852	64,42907	84,18767	71,71891	80,07581
8	86,49967	69,58494	54,58194	41,49067	64,06752	49,73632	59,07349
9	116,05925	100,03342	84,78380	70,31038	94,43104	78,86866	88,53055
10	113,51391	97,29438	81,98292	67,57954	90,27426	74,91978	83,22849

Os coeficientes das regressões, estimados pelos métodos TG1, TG2, TG3, e TG4 estão apresentados na Tabela 6 para C1 e na Tabela 7 para C2.

Para C1, foi observado que os coeficientes obtidos pelos métodos TG1 e TG2 apresentaram-se muito próximos, tendo o intercepto (b0) valores negativos e menores

que os dos métodos TG3 e TG4, indicando serem as médias iniciais dos valores genéticos preditos por TG1 e TG2 menores que os preditos por TG3 e TG4.

Tabela 6 – Coeficientes de regressão linear e quadráticos para os métodos TG1, TG2, TG3 e TG4 para a condição 1 (C1)

Table 6 – Linear and quadratic regression coefficients for the TG1, TG2, TG3 and TG4 methods, in the first condition (C1)

Rebanho <i>Herd</i>	CReg	TG1	TG2	TG3	TG4
1	b ₀	-4,33849859	-4,36209679	0,05546675	0,84009808
	b ₁	3,25042844	3,26733541	1,95878601	0,68203580
	b ₂	-0,12959941	-0,13165437	-0,15083078	0,09355971
2	b ₀	-6,00658464	-5,97346735	-0,97847295	-0,06299782
	b ₁	4,49615288	4,47289181	2,88704014	1,53839064
	b ₂	-0,24696435	-0,24418396	-0,22918954	0,02177785
3	b ₀	-5,26976442	-5,25861549	-1,47351813	-1,20392954
	b ₁	3,37722683	3,36813116	2,49780512	1,66619670
	b ₂	-0,13015366	-0,12899250	-0,21178365	-0,02884633
4	b ₀	-5,35014439	-5,31551075	-0,61391735	0,02928486
	b ₁	3,71530628	3,69143820	2,30943012	1,09184754
	b ₂	-0,16335689	-0,16046782	-0,18197839	0,06048224
5	b ₀	-5,27459764	-5,26339340	-0,49416459	0,40127811
	b ₁	3,46180296	3,45314479	1,92781043	0,61414719
	b ₂	-0,13394891	-0,13279296	-0,15559149	0,09825844
6	b ₀	-4,44837141	-4,42483044	-0,52507252	0,26289251
	b ₁	3,13867307	3,12341738	2,20080113	0,86989462
	b ₂	-0,09853220	-0,09693763	-0,16476427	0,09079683
7	b ₀	-5,49487782	-5,47198915	-0,77741683	0,20554681
	b ₁	3,92337132	3,90662026	2,76466966	1,32621694
	b ₂	-0,14912164	-0,14694600	-0,22033373	0,05328907
8	b ₀	-3,93565559	-3,90896964	0,67990702	1,36693299
	b ₁	3,55532670	3,53644991	2,28900218	0,98061305
	b ₂	-0,13719055	-0,13495733	-0,18883324	0,07373413
9	b ₀	-5,67713404	-5,63292170	-1,59356332	-0,40788722
	b ₁	4,17128038	4,14041376	3,31751323	1,70697451
	b ₂	-0,17440504	-0,17065263	-0,26685637	0,02927119
10	b ₀	-5,76799726	-5,76021194	-0,47520101	0,45799056
	b ₁	3,77230930	3,76747060	2,05725908	0,66623515
	b ₂	-0,13878365	-0,13831978	-0,16773266	0,10573889

Tabela 7 – Coeficientes de regressão linear e quadráticos para os métodos TG1, TG2, TG3 e TG4 para a condição 2 (C2)

Table 7 – Linear and quadratic regression coefficients for the TG1, TG2, TG3 and TG4 methods, in the second condition (C2)

Rebanho <i>Herd</i>	CReg	TG1	TG2	TG3	TG4
1	b ₀	0,12339	0,135281	-2,3E-07	0,26069
	b ₁	-0,39353	-0,402972	5,1179E-08	-0,32111
	b ₂	0,06293	0,064256	-7,342E-09	0,05774
3	b ₀	0,018891	0,024541	-7,176E-07	-0,059110
	b ₁	-0,096074	-0,097431	1,4508E-07	0,016157
4	b ₀	-0,985732	-1,068047	-7,82601E-06	-0,013538
	b ₁	0,701073	0,758251	4,30337E-06	-0,004790
	b ₂	-0,102551	-0,109307	-4,42596E-07	0,00070
6	b ₀	-0,554712	-0,565615	2,884E-07	-0,098460
	b ₁	0,121430	0,125114	-6,4E-08	0,028070
7	b ₀	-0,421865	-0,413742	-9,348E-07	0,002404
	b ₁	0,116556	0,115397	2,336E-07	-0,002045
8	b ₀	0,386301	0,403050	-4,2253E-06	0,016508
	b ₁	-0,363083	-0,375245	2,4856E-06	-0,034690
	b ₂	0,056490	0,058025	-2,6609E-07	0,005735
9	b ₀	-1,265163	-1,265927	-8,533E-08	-0,056557
	b ₁	0,560092	0,560907	3,6226E-08	0,047623
	b ₂	-0,061584	-0,061681	-1,114E-08	-0,008993
10	b ₀	0,010252	0,001578	-4,499E-06	0,079795
	b ₁	-0,280593	-0,278484	1,232E-06	-0,031125

Para a condição 2, foram observados valores muito baixos, tendendo a zero, para os coeficientes obtidos por meio do programa TG3.

Com o ajuste das regressões, foi possível a obtenção das estimativas das médias dos valores genéticos por ano, preditos para cada método em cada condição. O comportamento das curvas das tendências genéticas pode ser visto nas figuras 1 e 2 para C1 e 3 e 4 para C2.

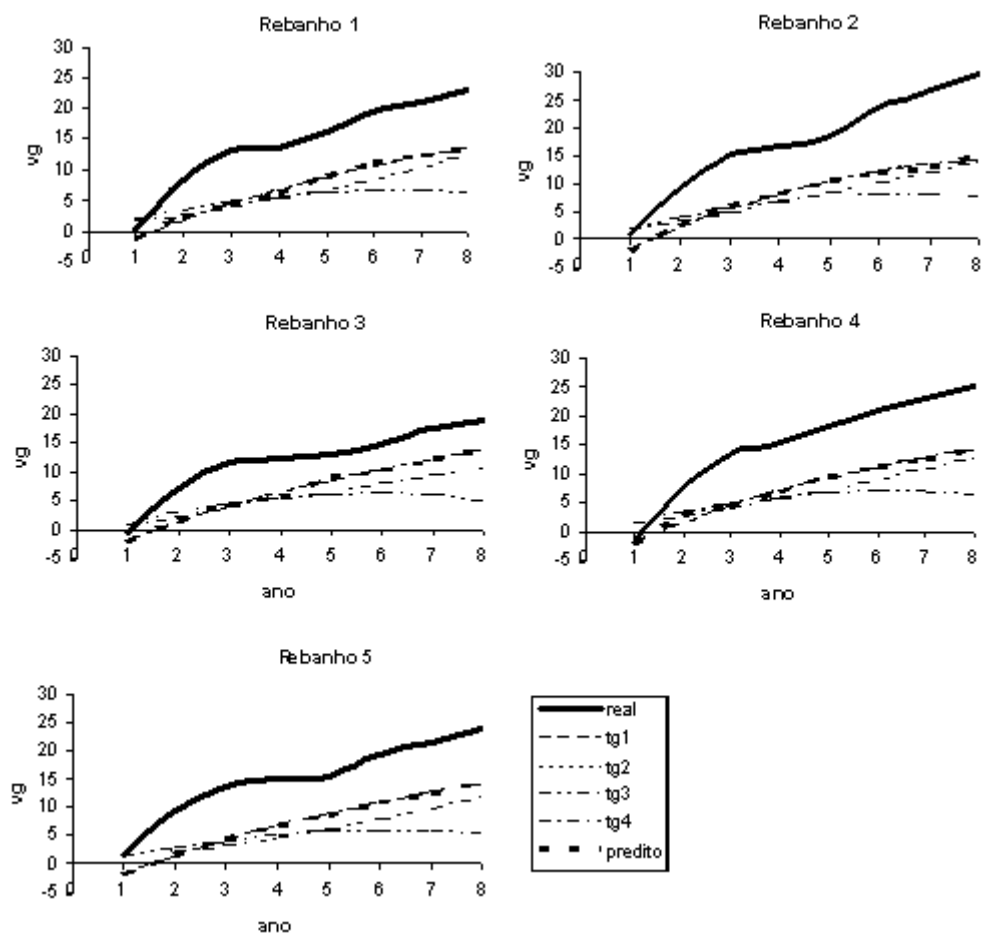


Figura 1 – Curvas de regressão das médias dos valores genéticos reais, preditos pela avaliação genética e pelos métodos TG1, TG2, TG3 e TG4, em função do ano de nascimento do animal, para os rebanhos de 1 a 5, na condição 1 (C1). Vg é o valor genético.

Figure 1 – Regression curves of the means of the real genetic values, predicted by the genetic evaluation and by the TG1, TG2, TG3 and TG4 methods as a function of the birth's year, for the herds 1 to 5, in the first condition (C1). Vg is the genetic value.

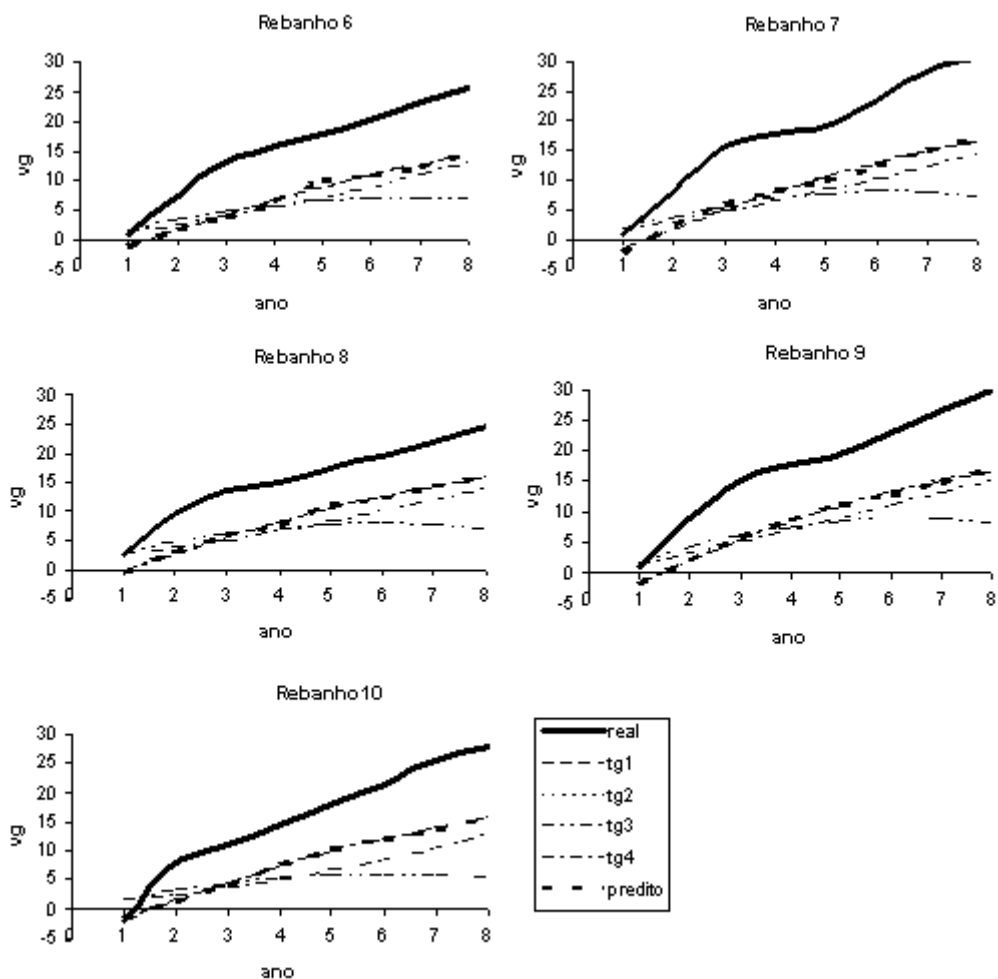


Figura 2 – Curvas de regressão das médias dos valores genéticos reais, preditos pela avaliação genética e pelos métodos TG1, TG2, TG3 e TG4, em função do ano de nascimento do animal, para os rebanhos de 6 a 10, na condição 1 (C1). Vg é o valor genético.

Figure 2 – Regression curves of the means of the real genetic values, predicted by genetic evaluation and by the TG1, TG2, TG3 and TG4 methods as a function of the birth's year, for the herds 6 to 10, in the first condition (C1). Vg is the genetic value.

Na condição 1, para todos os rebanhos, ficou notada a superioridade dos valores médios estimados por meio do método TG1, em relação aos estimados por TG3 e TG4, e os valores de TG2 muito próximos aos de TG1 e praticamente não-distinguíveis entre si e entre as médias dos valores genéticos preditos pela avaliação.

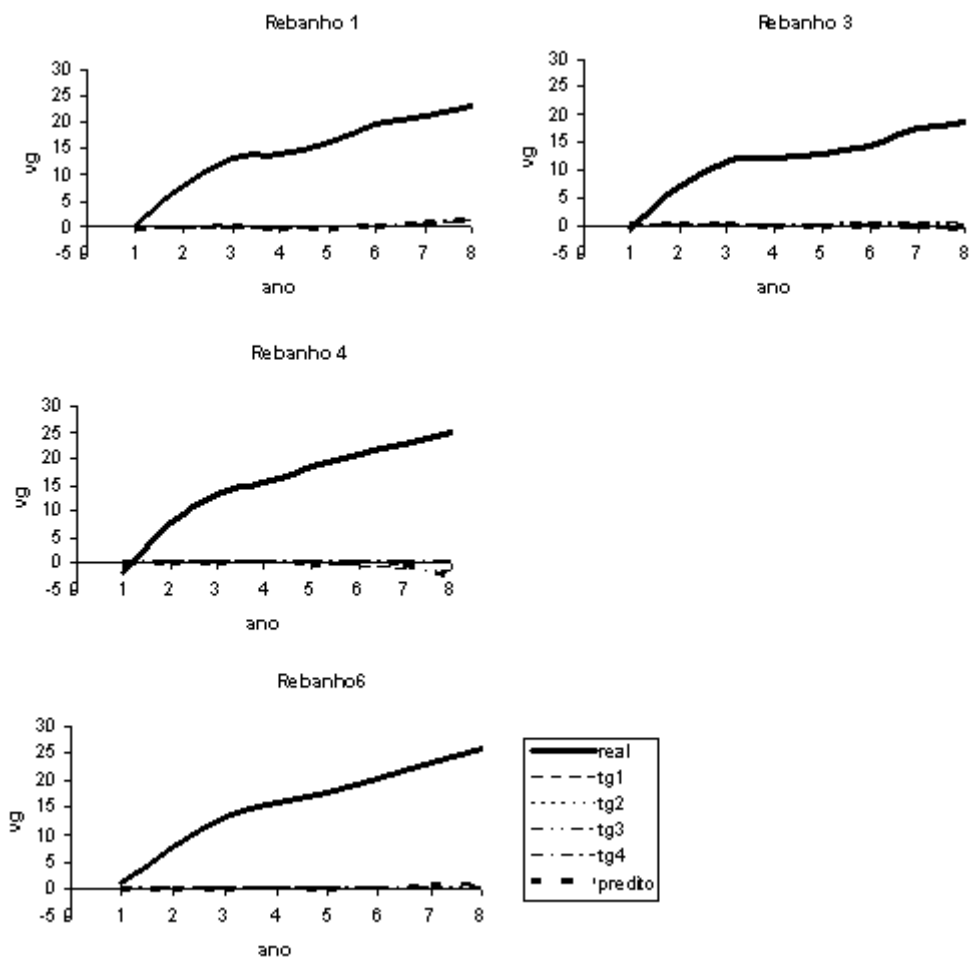


Figura 3 – Curvas de regressão das médias dos valores genéticos reais, preditos pela avaliação genética e pelos métodos TG1, TG2, TG3 e TG4, em função do ano de nascimento do animal, para os rebanhos de 1 a 6, na condição 2 (C2). Vg é o valor genético.

Figure 3 – Regression curves of the means of the real genetic values, predicted by genetic evaluation and by the TG1, TG2, TG3 and TG4 methods as a function of the birth's year, for the herds 1 to 6, in the second condition (C2). Vg is the genetic value.

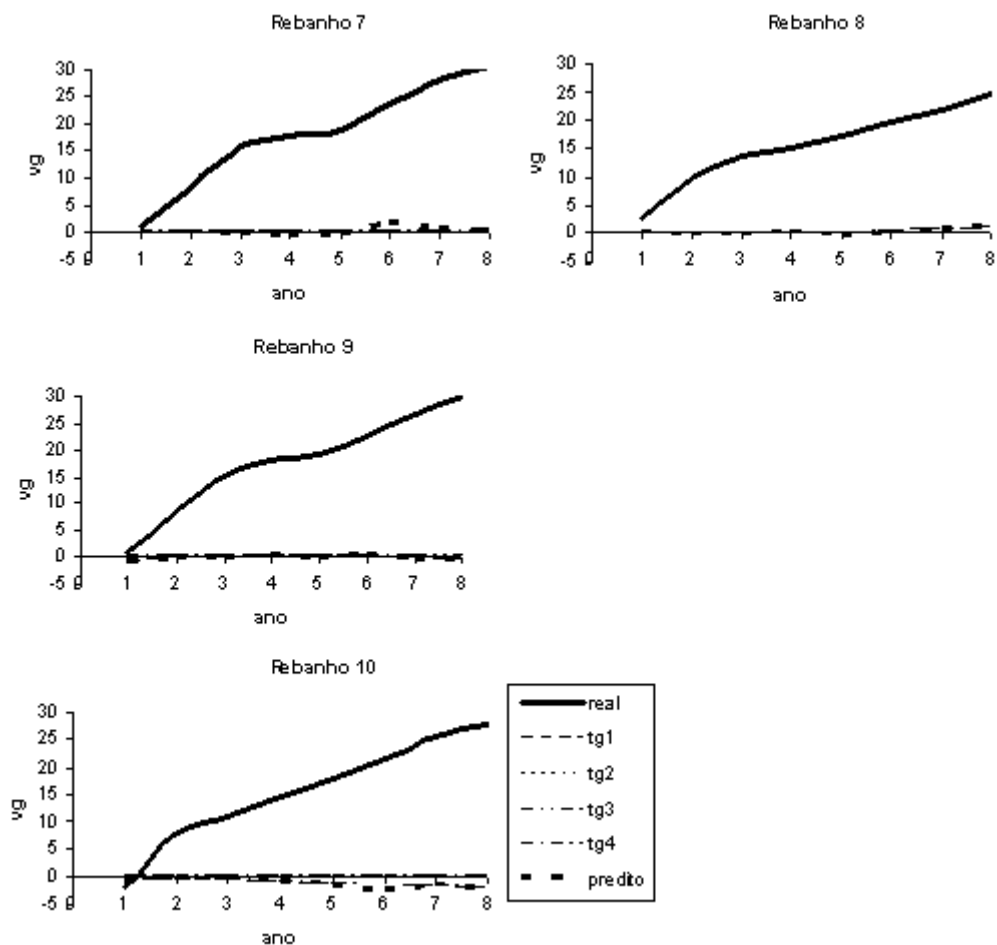


Figura 4 – Curvas de regressão das médias dos valores genéticos reais, preditos pela avaliação genética e pelos métodos TG1, TG2, TG3 e TG4, em função do ano de nascimento do animal, para os rebanhos de 7 a 10, na condição 2 (C2). V_g é o valor genético.

Figure 4 – Regression curves of the means of the real genetic values, predicted by genetic evaluation and by the TG1, TG2, TG3 and TG4 methods as a function of the birth's year, for the herds 7 to 10, in the second condition (C2). V_g is the genetic value.

Para a condição 2, foi evidente a inferioridade das médias dos valores genéticos obtidos pelos métodos TG1, TG2, TG3 e TG4 e pela avaliação genética em relação às médias dos valores genéticos reais, que indica ocorrência de subestimação pela ausência de uma estrutura robusta de parentesco. Não foram mostradas curvas para os rebanhos 2 e 5 porque não houve efeito do ano de nascimento do animal sobre as médias dos valores genéticos.

Conforme Schaeffer (1996), as tendências genéticas são facilmente calculadas pela média dos valores genéticos estimados de grupos de animais por ano, no entanto,

seu cálculo deve ser precisamente definido. Deve-se tomar cuidado com estimativa de herdabilidade assumida para as equações de modelos mistos (MME); se ela é maior ou menor que a herdabilidade verdadeira, a tendência genética pode ser superestimada ou subestimada. O mesmo autor defende ainda que o processo de seleção deve ser levado em conta em procedimentos que avaliem o desempenho genético dos animais, pois esse causa a redução da variabilidade genética pelo aumento do desequilíbrio gamético.

O teste de aderência entre as curvas de regressão ajustadas para o valor genético predito e para os valores genéticos estimados pelos métodos TG1, TG2, TG3, TG4, forneceu baixos coeficientes de determinação, crescentes na ordem TG3, TG4, TG2, TG1, para as duas condições, com médias 0,13715; 0,36587; 0,42997; 0,42998, respectivamente, para a condição 1 e -0,00290; -0,00212; 0,00245; 0,00246, respectivamente para a condição 2, e indicam serem os métodos TG1 e TG2 os que mais se aproximam ao comportamento dos valores genéticos preditos pela avaliação, tanto na C1 como na C2.

Um alto valor de R^2 não necessariamente implica que o modelo ajustado se presta para fazer inferências precisas, pois apesar de um alto valor de R^2 , o quadrado médio do resíduo ainda pode ser grande, indicando ajuste duvidoso por ser alta a variância não-explicada pelo modelo.

A mensuração da distância entre os valores genéticos reais e os estimados, pelos métodos TG1, TG2, TG3, TG4, indicou ser o método TG3 o que fornece valores mais distantes dos reais, pois teve maior soma de quadrados residual, sendo a ordem de aproximação TG1, TG2, TG4, TG3, com médias 1.544.777,69; 1.544.783,94; 1.826.359,41; 2.326.534,66, respectivamente, para a condição 1. Na condição 2, o método TG1 foi o que apresentou maior soma de quadrados residual, portanto, o que mais se distanciou dos valores genéticos reais. A ordem de aproximação foi TG4, TG3, TG2, TG1, com médias 4.021.510,91; 4.033.257,81; 4.127.355,59; 4.128.306,63. Isso indica que

quando o objetivo é fazer o estudo da tendência genética estimada versus a tendência real, em determinados grupos de animais pertencentes a um rebanho ou em rebanhos pertencentes a uma associação de raça, a melhor opção é considerar no modelo a existência de heterogeneidade de variâncias e a estrutura de parentesco entre os indivíduos, como foi considerado no método TG4. No entanto, quando o objetivo é avaliar a tendência genética em rebanhos como um todo, ou na raça como um todo, o melhor é fazer a regressão simples dos valores genéticos ao longo dos anos, visto que a inclusão de informações não foi capaz de dar uma melhor explicação ao modelo. Em todos os casos, é importante verificar se o valor genético sofre efeito linear ou quadrático do ano.

Euclides Filho et al. (2000), ao estudar a tendência genética em animais da raça Gir, afirmaram que as estimativas de tendências genéticas, obtidas em uma população criada em uma grande diversidade de ambientes e nos mais diversos sistemas de produção, como é o caso de gado de corte no Brasil, são o resultado médio de valores que podem encontrar-se distribuídos, alguns próximos a zero, enquanto outros são positivos ou negativos. Se isso, por um lado, permite avaliar a situação da raça como um todo, por outro, impossibilita a identificação dos rebanhos que possam realizar os maiores progressos. Os autores defendem ainda que programas estruturados com base na demanda e no potencial de diferentes regiões podem possibilitar progressos genéticos mais eficientes tanto do ponto de vista biológico quanto do econômico.

Conclusões

Na maioria dos rebanhos, observou-se efeito quadrático do ano de nascimento sobre os valores genéticos dos animais, com aumento dos valores genéticos e redução dos ganhos genéticos.

Houve a subestimação dos valores genéticos quando a avaliação não considerou a estrutura total de parentesco existente.

O método para a estimação da tendência genética a ser utilizado quando o interesse está em se aproximar do valor genético predito pela avaliação e quando se trata da população como um todo é o TG1.

Quando o interesse é a aproximação ao valor genético real, em subpopulações, deve-se considerar a estrutura de parentesco e a existência de heterogeneidade de (co)variância, como foi considerado no método TG4.

Literatura Citada

- BOLDMAN, K.G.; KRIESE, L.A.; VAN VLECK, L.D. et al. **A manual fo use of MTDFREML. A set of programs to obtain estimates of variances and covariances[Draft]**. USDA-ARS, Lincoln,NE, 1995.
- COWLES, K.; BEST, N.; VINES, K. **Convergence diagnosis andoutput analysis**. Cambridge: MRC Biostatistics Unit, UK.Version 0.40. 1995.
- EUCLIDES FILHO, K.; NOBRE, P.R.C.; ROSA, A.N. Tendência genética na raça Guzerá. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 34. 1997, Juiz de Fora. **Anais...** Juiz de Fora: Sociedade Brasileira de Zootecnia, 1997. p.175.
- EUCLIDES FILHO, K.; SILVA, L. O. C.; ALVES, R. G. O. et al. Tendência genética na raça Gir. **Pesquisa Agropecuária Brasileira**. v.35, n.4, p.787-791, 2000.
- FERNANDES, H.D.; FERREIRA G.B.B.; RORATO, P.R.N. Tendências e parâmetros genéticos para características pré-desmama em bovinos da raça Charolês criados no Rio Grande do Sul. **Revista Brasileira de Zootecnia**, v.31, n.1, p.321-330, 2002.
- FERRAZ FILHO, P.B.; BIANCHINI SOBRINHO, E.; SILVA, L.O.C. et al. Tendência genética em pesos de bovinos da raça Nelore mocha no Brasil. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 34, 1997, Juiz de Fora. **Anais...** Juiz de Fora: Sociedade Brasileira de Zootecnia, 1997.
- FERRAZ FILHO, P.B.; RAMOS, A.A.; SILVA, L.O.C. et al. Tendência genética dos efeitos direto e materno sobre os pesos à desmama e pós-desmama de bovinos da raça Tabapuã no Brasil. **Revista Brasileira de Zootecnia**, v.31, n.2, p.635-640, 2002.
- 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 4**. Oxford: University Press, 1992. p.625-631.

- HEIDELBERGER, P.; WELCH, P.D. Simulation run length control in the presence of an initial transient. **Operations Research**, Landing, v.31, n.6, p.1109-1144, 1983.
- HENDERSON, C. R. **Applications of linear models in animal breeding**. University of Guelph, Ontario, Canadá, 1984.
- HOLANDA, M.C.R.; BARBOSA, S.B.P.; RIBEIRO, A.C. et al. Tendências genéticas para crescimento em bovinos Nelore em Pernambuco, Brasil. **Arch. Zootec.**, v.53, p.185-194, 2004.
- KENNEDY, B. W.; SCHAEFFER, L. R.; SORENSEN, D. A. Genetic properties of animal models. **Journal of Dairy Science**, v.71, p.17-26, 1988 supplement.
- LÔBO, R.B.; OLIVEIRA, H.N.; BEZERRA, L.F. et al. Estimativa de componentes de (co)variância herdabilidade para o peso aos 120 dias de idade na raça Nelore usando estatística bayesiana. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 34, 1997, Juiz de Fora – MG. **Anais...** Juiz de Fora: SBZ, 1997. p. 186-189.
- MARTINS, E.N. Uso de modelos mistos na avaliação genética animal. Viçosa: UFV, 121p. 1997.
- MELLO, S.P. **Tendência genética para pesos em um rebanho da raça Canchim**. Jaboticabal: Universidade Estadual Paulista, 1999. 78p. Dissertação (Mestrado em Zootecnia) – Universidade Estadual Paulista, 1999.
- OLIVEIRA, C.A.L. **Estimadores e preditores em um modelo linear misto, utilizando-se mínimos quadrados ordinários e mínimos quadrados generalizados e a melhor predição linear não viesada**. Maringá: Universidade Estadual de Maringá, 1998. 21p. Monografia (Especialização em Estatística) – Universidade Estadual de Maringá, 1998.
- OLIVEIRA, C.A.L. **Efeito direto e materno em características de crescimento em rebanhos de bovinos de corte sob seleção**. 2006. 85p. Tese (Doutorado em Zootecnia) – Universidade Estadual de Maringá, 2006.
- SCHAEFFER, L.R. Animal Models: Why, How and When to Use Them. In: Simpósio Nacional de Melhoramento animal. Sociedade Brasileira de Melhoramento Animal, 1, 1996, Ribeirão. Preto, SP. **Anais...** Ribeirão Preto: Sociedade Brasileira de Melhoramento Animal, 1996. p.21-40.
- SILVA, L.O.C; FILHO, K.E.; NOBRE, P.R.C. et al. Tendências genéticas na raça Nelore no Brasil. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 34, 1997, Juiz de Fora. **Anais...** Juiz de Fora: Sociedade Brasileira de Zootecnia, 1997a. p.175-176.
- SMITH, C. Rates of genetic change in farm livestock. **Research Development Agricultural**, v.1, n.2, p.79-85; **Animal Breeding Abstracts**, v.53, n.9, p.699, 1985.
- VAN TASSEL, C.P.; VAN VLECK, L.D. **A manual for use of MTGSAM. A set of FORTRAN programs to apply Gibbs sampling to animal models for variance component estimation (DRAFT)**. Lincoln: Department of Agriculture Research Service, 1995. 86p.
- ZOLLINGER, W.A.; NIELSEN, M.K. An evaluation of bias in estimated breeding values for weaning weight in Angus beef cattle field records. I. Estimates of within herd genetic trend. **Journal of Animal Science**, v.58, p.545-549, 1984.

V – CONCLUSÕES GERAIS

O modelo de regressão múltipla dos componentes de (co)variância em função das gerações sob seleção pode ser implementado em procedimentos bayesianos para estimação de componentes de (co)variância quando se admite heterocedasticidade ao longo do tempo, por efeito da seleção.

Na maioria dos rebanhos observou-se efeito quadrático do ano de nascimento sobre os valores genéticos dos animais, com aumento dos valores genéticos e redução dos ganhos genéticos. Houve a subestimação dos valores genéticos quando a avaliação não considerou a estrutura total de parentesco existente.

O método para a estimação da tendência genética a ser utilizado quando o interesse está em se aproximar do valor genético predito pela avaliação e quando se trata da população, como um todo, é o TG1.

Quando o interesse é a aproximação ao valor genético real, em subpopulações, deve-se considerar a estrutura de parentesco e a existência de heterogeneidade de (co)variância, conforme foi considerado no método TG4.

VI – APÊNDICE

Tabela 1A – Componentes de variância genética aditiva estimados por ano para P550

Table 1A – Components of additive genetic variances estimated by year for P550

Componente <i>Component</i>	Rebanhos <i>Herds</i>									
	1	2	3	4	5	6	7	8	9	10
$\sigma_{a_{1-13}}^2$	91,33903	106,18484	86,60608	89,84806	87,86214	85,47890	103,02475	86,49967	116,05925	113,51391
$\sigma_{a_{14}}^2$	84,90835	102,11155	80,74418	82,96075	80,02488	80,21008	97,85220	80,31046	110,76962	106,53682
$\sigma_{a_{15}}^2$	78,88010	98,93034	75,22620	76,91037	72,64291	75,40450	93,34388	74,64462	105,18190	99,53407
$\sigma_{a_{16}}^2$	73,25426	96,64122	70,05212	71,69693	65,71625	71,06213	89,49979	69,50215	99,29606	92,50564
$\sigma_{a_{17}}^2$	68,03086	95,24416	65,22195	67,32043	59,24487	67,18298	86,31992	64,88306	93,11211	85,45155
$\sigma_{a_{18}}^2$	63,20989	94,73919	60,73570	63,78086	53,22879	63,76707	83,80428	60,78735	86,63005	78,37178
$\sigma_{a_{19}}^2$	58,79134	95,12629	56,59336	61,07822	47,66800	60,81439	81,95287	57,21501	79,84989	71,26635
$\sigma_{a_{20}}^2$	54,77522	96,40546	52,79494	59,21252	42,56250	58,32493	80,76569	54,16604	72,77161	64,13525
$\sigma_{a_{21}}^2$	51,16151	98,57672	49,34042	58,18377	37,91230	56,29869	80,24273	51,64045	65,39523	56,97847
$\sigma_{a_{22}}^2$	47,95024	101,64005	46,22982	57,99194	33,71740	54,73569	80,38400	49,63824	57,72073	49,79604
$\sigma_{a_{23}}^2$	45,14140	105,59547	43,46313	58,63705	29,97778	53,63591	81,18950	48,15940	49,74813	42,58793
$\sigma_{a_{24}}^2$	42,73498	110,44295	41,04036	60,11910	26,69346	52,99936	82,65923	47,20394	41,47742	35,35415
$\sigma_{a_{25}}^2$	40,73099	116,18251	38,96149	62,43808	23,86443	52,82603	84,79318	46,77185	32,90860	28,09471
$\sigma_{a_{26}}^2$	39,12942	122,81416	37,22654	65,59400	21,49070	53,11592	87,59136	46,86314	24,04167	20,80959
$\sigma_{a_{27}}^2$	37,93028	130,33788	35,83550	69,58686	19,57226	53,86905	91,05377	47,47780	14,87663	13,49881
$\sigma_{a_{28}}^2$	37,13356	138,75366	34,78837	74,41665	18,10911	55,08541	95,18040	48,61583	5,41348	6,16236

Tabela 2A – Componentes de covariância genética aditiva estimados por ano para P550, para os rebanho de 1 a 5

Table 2A – Components of additive genetic covariance estimated by year for P550, for herds 1 to 5

Componente <i>Component</i>	Rebanhos <i>Herds</i>				
	1	2	3	4	5
$\sigma_{a_{1-13}a_{1-13}}$	91,33903	106,18484	86,60608	89,84806	87,86214
$\sigma_{a_{14}a_{15}}$	69,84147	87,64512	66,35640	69,36311	64,04747
$\sigma_{a_{14}a_{16}}$	56,54684	74,36935	53,66952	57,19348	49,80622
$\sigma_{a_{14}a_{17}}$	45,02446	62,28428	42,68353	46,45187	37,30114
$\sigma_{a_{14}a_{18}}$	35,27433	51,38988	33,39844	37,13828	26,53221
$\sigma_{a_{15}a_{16}}$	64,50056	84,76817	61,51760	64,00409	57,33769
$\sigma_{a_{15}a_{17}}$	51,89328	71,79668	49,50990	52,52583	43,76862
$\sigma_{a_{15}a_{18}}$	41,05824	60,01588	39,20309	42,47557	31,93572
$\sigma_{a_{16}a_{17}}$	59,56208	82,78332	57,02271	59,48201	51,08320
$\sigma_{a_{16}a_{18}}$	47,64214	70,11610	45,69419	48,69511	38,18632
$\sigma_{a_{17}a_{18}}$	55,02602	81,69053	52,87172	55,79686	45,28401
$\sigma_{a_{19}a_{20}}$	47,16118	82,18119	45,60150	50,93738	35,05150
$\sigma_{a_{19}a_{21}}$	37,30326	70,42677	36,31054	42,22456	24,17117
$\sigma_{a_{19}a_{22}}$	29,21760	59,86304	28,72047	34,93975	15,02700
$\sigma_{a_{19}a_{23}}$	22,90419	50,48998	22,83130	29,08296	7,61899
$\sigma_{a_{20}a_{21}}$	43,83240	83,76464	42,48226	49,76305	30,61819
$\sigma_{a_{20}a_{22}}$	34,66183	72,31449	33,87048	41,74158	20,41004
$\sigma_{a_{20}a_{23}}$	27,26351	62,05502	26,95960	35,14814	11,93805
$\sigma_{a_{21}a_{22}}$	40,90604	86,24017	39,70693	49,42565	26,64017
$\sigma_{a_{21}a_{23}}$	32,42282	75,09428	31,77433	42,09554	17,10420
$\sigma_{a_{22}a_{22}}$	38,38211	89,60777	37,27551	49,92518	23,11744
$\sigma_{a_{24}a_{25}}$	34,54153	99,01920	33,44441	53,43506	17,43788
$\sigma_{a_{24}a_{26}}$	28,12034	88,78612	27,54937	48,17904	9,91846
$\sigma_{a_{24}a_{27}}$	23,47139	79,74372	23,35522	44,35103	4,13520
$\sigma_{a_{24}a_{28}}$	20,59469	71,89201	20,86197	41,95104	0,08810
$\sigma_{a_{25}a_{26}}$	33,22488	105,06303	32,04473	56,44541	15,28103
$\sigma_{a_{25}a_{27}}$	27,49103	95,13422	26,82887	51,88074	8,43379
$\sigma_{a_{25}a_{28}}$	23,52943	86,39609	23,31391	48,74410	3,32272
$\sigma_{a_{26}a_{27}}$	32,31066	111,99893	30,98897	60,29269	13,57948
$\sigma_{a_{26}a_{28}}$	27,26415	102,37440	26,45229	56,41938	7,40443
$\sigma_{a_{27}a_{28}}$	31,79886	119,82692	30,27711	64,97691	12,33322

Tabela 3A – Componentes de covariância genética aditiva estimados por ano para P550, para os rebanho de 6 a 10

Table 3A – Components of additive genetic covariance estimated by year for P550, for herds 6 to 10

Componente <i>Component</i>	Rebanhos <i>Herds</i>				
	6	7	8	9	10
$\sigma_{a_{1-13}a_{1-13}}$	85,47890	103,02475	86,49967	116,05925	113,51391
$\sigma_{a_{14}a_{15}}$	67,00462	84,18767	64,06752	94,43104	90,27426
$\sigma_{a_{14}a_{16}}$	55,37115	71,71891	49,73632	78,86866	74,91978
$\sigma_{a_{14}a_{17}}$	45,30970	60,44593	37,31686	64,08249	60,47338
$\sigma_{a_{14}a_{18}}$	36,82024	50,36872	26,80913	50,07252	46,93505
$\sigma_{a_{15}a_{16}}$	62,80071	80,07581	59,07349	88,53055	83,22849
$\sigma_{a_{15}a_{17}}$	51,76893	68,00352	45,41409	72,65541	67,83098
$\sigma_{a_{15}a_{18}}$	42,30916	57,12701	33,66643	57,55648	53,34155
$\sigma_{a_{16}a_{17}}$	59,06003	76,62819	54,60283	82,33195	76,15704
$\sigma_{a_{16}a_{18}}$	48,62994	64,95237	41,61523	66,14406	60,71650
$\sigma_{a_{17}a_{18}}$	55,78258	73,84479	50,65554	75,83525	69,05991
$\sigma_{a_{19}a_{20}}$	50,61735	70,27068	44,33109	61,94751	54,78867
$\sigma_{a_{19}a_{21}}$	41,99231	59,78426	33,35891	44,82134	39,21907
$\sigma_{a_{19}a_{22}}$	34,93928	50,49363	24,29846	28,47138	24,55754
$\sigma_{a_{19}a_{23}}$	29,45826	42,39877	17,14974	12,89762	10,80409
$\sigma_{a_{20}a_{21}}$	48,72957	69,47997	41,95393	54,55648	47,61454
$\sigma_{a_{20}a_{22}}$	40,70622	59,39001	31,65355	37,11755	32,00191
$\sigma_{a_{20}a_{23}}$	34,25488	50,49585	23,26490	20,45483	17,29736
$\sigma_{a_{21}a_{22}}$	47,30503	69,35348	40,10014	46,86734	40,41475
$\sigma_{a_{21}a_{23}}$	39,88336	59,66000	30,47157	29,11565	24,75910
$\sigma_{a_{22}a_{22}}$	46,34370	69,89121	38,76973	38,88008	33,18929
$\sigma_{a_{24}a_{25}}$	45,81074	72,95937	37,67904	22,01125	18,66135
$\sigma_{a_{24}a_{26}}$	40,19413	64,45531	30,06587	3,32129	2,87663
$\sigma_{a_{24}a_{27}}$	36,14953	57,14701	24,36443	-14,59246	-12,00002
$\sigma_{a_{24}a_{28}}$	33,67693	51,03449	20,57473	-31,73001	-25,96860
$\sigma_{a_{25}a_{26}}$	46,23910	75,48980	37,91875	13,12967	11,35888
$\sigma_{a_{25}a_{27}}$	41,22418	67,38219	30,97738	-5,87305	-4,46887
$\sigma_{a_{25}a_{28}}$	37,78126	60,47037	25,94775	-24,09956	-19,38854
$\sigma_{a_{26}a_{27}}$	47,13069	78,68445	38,68184	3,94998	4,03074
$\sigma_{a_{26}a_{28}}$	42,71745	70,97331	32,41228	-15,36549	-11,84003
$\sigma_{a_{27}a_{28}}$	48,48550	82,54333	39,96830	-5,52782	-3,32307

