

UNIVERSIDADE ESTADUAL PAULISTA "JULIO DE MESQUITA FILHO"  
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS  
CÂMPUS DE JABOTICABAL

CRITÉRIOS DE SELEÇÃO PARA INCREMENTO DE  
UNIFORMIDADE DE PRODUÇÃO EM BOVINOS DE CORTE

Haroldo Henrique de Rezende Neves  
Zootecnista

JABOTICABAL – SÃO PAULO – BRASIL  
Fevereiro de 2010

UNIVERSIDADE ESTADUAL PAULISTA "JULIO DE MESQUITA FILHO"  
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS  
CÂMPUS DE JABOTICABAL

CRITÉRIOS DE SELEÇÃO PARA INCREMENTO DE  
UNIFORMIDADE DE PRODUÇÃO EM BOVINOS DE CORTE

Haroldo Henrique de Rezende Neves

Orientador: Prof<sup>a</sup>. Dr<sup>a</sup>. Sandra Aidar de Queiroz

Co-orientador: Dr. Roberto Carneiro

Dissertação apresentada à Faculdade de Ciências Agrárias e Veterinárias – Unesp, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Mestre em Genética e Melhoramento Animal.

JABOTICABAL – SÃO PAULO – BRASIL

Fevereiro de 2010

N514c Neves, Haroldo Henrique de Rezende  
Critérios de seleção para incremento de uniformidade de  
produção em bovinos de corte / Haroldo Henrique de Rezende Neves.  
-- Jaboticabal, 2010  
iv, 96 f. : il. ; 28 cm

Dissertação (mestrado) - Universidade Estadual Paulista,  
Faculdade de Ciências Agrárias e Veterinárias, 2010  
Orientadora: Sandra Aidar de Queiroz  
Banca examinadora: Roberto Carneiro, Henrique Nunes de  
Oliveira, Vanerlei Mozaquatro Roso  
Bibliografia

1. Melhoramento Genético. 2. Bovinos de corte. 3. Variância  
residual. I. Título. II. Jaboticabal-Faculdade de Ciências Agrárias e  
Veterinárias.

CDU 636.2:636.082

Ficha catalográfica elaborada pela Seção Técnica de Aquisição e Tratamento da Informação – Serviço  
Técnico de Biblioteca e Documentação - UNESP, Câmpus de Jaboticabal.



UNIVERSIDADE ESTADUAL PAULISTA

CAMPUS DE JABOTICABAL

FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS DE JABOTICABAL

**CERTIFICADO DE APROVAÇÃO**

**TÍTULO:** CRITÉRIOS DE SELEÇÃO PARA INCREMENTO DE UNIFORMIDADE DE PRODUÇÃO EM BOVINOS DE CORTE.

**AUTOR:** HAROLDO HENRIQUE DE REZENDE NEVES  
**CO-ORIENTADOR:** Prof. Dr. ROBERTO CARVALHEIRO

Aprovado como parte das exigências para obtenção do Título de MESTRE em GENÉTICA E MELHORAMENTO ANIMAL , pela Comissão Examinadora:

Prof. Dr. ROBERTO CARVALHEIRO  
Melhoramento Genético Animal, GenSys Consultores Associados S/C Ltda, Jaboticabal/SP

Prof. Dr. VANERLEI MOZAQUATRO ROSO  
Melhoramento Genético Animal, GenSys Consultores Associados S/C Ltda, Porto Alegre/RS

Prof. Dr. HENRIQUE NUNES DE OLIVEIRA  
Departamento de Melhor e Nutri Animal / Faculdade de Medicina Veterinária e Zootecnia de Botucatu

Data da realização: 01 de fevereiro de 2010.

## **DADOS CURRICULARES DO AUTOR**

**HAROLDO HENRIQUE DE REZENDE NEVES** – nasceu em 15 de junho de 1984, no município de Araxá – MG, filho de Cecília Maria de Rezende Neves e Antonino Neves de Resende. Iniciou o curso de Zootecnia na FCAV/Unesp – Jaboticabal em fevereiro de 2003 e concluiu em dezembro de 2007. Durante a graduação também foi bolsista do Programa de Educação Tutorial (PET) e professor do curso pré-vestibular Ativo. Defendeu o trabalho de graduação intitulado “Aplicações de acasalamento dirigido e sêmen sexado em programas de melhoramento de zebuínos de corte”, sob orientação da Prof<sup>a</sup>. Dr<sup>a</sup>. Sandra Aidar de Queiroz. Iniciou o curso de Mestrado em Genética e Melhoramento Animal na FCAV/Unesp em março de 2008, tendo sido bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior e, posteriormente, da Fundação de Amparo à Pesquisa do Estado de São Paulo. Submeteu-se ao exame de defesa de dissertação para obtenção do título de Mestre em Genética e Melhoramento Animal em 01 de fevereiro de 2010, sob orientação da Prof<sup>a</sup>. Dr<sup>a</sup>. Sandra Aidar de Queiroz e co-orientação do Dr. Roberto Carvalheiro.

“O correr da vida embrulha tudo.  
A vida é assim: esquenta e esfria,  
aperta e daí afrouxa,  
sossega e depois desinquieta.  
O que ela quer da gente é coragem.”

**Guimarães Rosa – Grande Sertão: Veredas**

### **"Lembrete**

Se procurar bem você acaba encontrando.  
não a explicação (**duvidosa**) da vida,  
mas a poesia (**inexplicável**) da vida."

***Carlos Drummond de Andrade***

À minha saudosa mãe Cecília, por todo amor, dedicação, coragem, fé,  
ensinamentos...você sempre será parte importante desta minha caminhada.  
("Mãe, na sua graça, é eternidade.")

Ao meu pai Antonino, pelo incentivo, por todos os exemplos de trabalho e integridade e  
pela oportunidade de, desde cedo, criar raízes com a terra.

Ao meu irmão Aurélio, pela amizade, apoio, companheirismo.

À toda minha família...vocês são demais!

**Dedico**

## AGRADECIMENTOS

A Deus, pelo milagre da vida...pela oportunidade de ter vivido 9.366 dias até o momento em que estas linhas foram escritas e pela força para seguir nesta caminhada.

À professora Sandra Aidar de Queiroz pela orientação, amizade, incentivo e apoio. Minha sincera gratidão e admiração.

Ao amigo Roberto Carvalheiro pela orientação, disposição, conselhos, apoio e incentivo e por toda contribuição na minha formação ao longo destes últimos cinco anos, minha profunda amizade, gratidão e admiração.

Aos membros das bancas examinadoras de qualificação e de defesa, Dr. Danísio Prado Munari, Dr. Henrique Nunes de Oliveira e Dr. Vanerlei Mozaquatro Roso, pela disponibilidade, pelas valiosas sugestões e contribuições para a melhoria deste trabalho.

À Gensys Consultores Associados pela disponibilização dos dados utilizados neste trabalho.

À Faculdade de Ciências Agrárias e Veterinárias - Unesp Jaboticabal pela oportunidade de formação e por estar ligada a alguns bons momentos de minha vida.

À FAPESP e CAPES pela concessão de bolsa de estudos durante períodos distintos do Mestrado.

Ao Programa de Pós-Graduação em Genética e Melhoramento Animal, pelo curso oferecido e pelo apoio imprescindível em diversas etapas deste trabalho.

A todos os professores do Programa de Pós-Graduação em Genética e Melhoramento Animal, pela contribuição à minha formação acadêmica.

A todos os colegas de Pós-graduação, em especial: Diego Barrozo, Milene, Aila, Leonardo, Biro, Eduardo, Daniel Gordo, Diego Guidolin, Dani, Salvador, Rodrigo, Juana, Francisco, Ana Paula, Raúl, Severino, Tomás...

A todos funcionários e professores da FCAV, em especial à Prof<sup>a</sup>. Maria Imaculada, tutora do PET Zootecnia, e ao Prof. João Ademir, "culpado" pelo início de meus estudos na área de Melhoramento Animal.

Aos amigos da República Mata Bixera, em ordem hierárquica: Gaúxo, Bernento, Lírou, Cervo, Pedala, Teta, Mininão, Sofredô, Gozo, Rombo, Biruleiby, Xurros, Muralha, Delegado, pelo excelente convívio e pelos momentos surreais que estas retinas tão fatigadas puderam presenciar. À nossa governanta, Dona Madalena, pela dedicação, cuidado e paciência. Aos veteranos mais presentes, Finkado, Batata, Grico e Pinico e a todos que sempre prezaram pela manutenção das tradições e bons costumes da república ao longo das últimas duas décadas. E, claro, à Luna, nossa inesquecível representante dos *Canis familiaris*.

Aos grandes amigos da época de graduação cuja presença é motivo de alegria e muitas risadas: Mama-ki, Bodão, Furunku, Xoco, Vurto, Passivo, Bago, Tansso, Mala...

Aos amigos da jurássica época de bixo e dos tempos de República Entrometheu, em especial Tumor, Bicudo, Fabiano, Koringa, Futum, Tiririca, Boto e Bigato.

Aos amigos da XLVII Turma de Zootecnia da FCAV, ou simplesmente ZOO 2003, (por questão de brevidade, a relação completa encontra-se na seção “Agradecimentos” de meu trabalho de graduação), pela convivência e companheirismo ao longo destes sete anos.

À Maíra, pelo carinho, companhia e pela alegria que tem trazido leveza a esses dias.

## SUMÁRIO

CAPÍTULO 1. CONSIDERAÇÕES GERAIS .....	1
1.1. Introdução.....	1
1.2. Modelos considerando heterogeneidade genética sobre a variância residual.....	3
1.3. Evidências de heterogeneidade genética sobre a variância residual .....	5
1.4. Medidas de dispersão e uso de estimadores alternativos às DEPs para seleção contemplando maior uniformidade.....	8
1.5. Objetivos.....	11
1.6. Referências.....	12
CAPÍTULO 2. SELEÇÃO CONTEMPLANDO UNIFORMIDADE NO GANHO DE PESO DO NASCIMENTO À DESMAMA EM BOVINOS NELORE: ESTUDOS PRELIMINARES .....	16
2.1. INTRODUÇÃO.....	17
2.2. MATERIAL E MÉTODOS .....	19
Estimadores de dispersão residual para incremento de uniformidade considerando efeito aditivo de touro (Modelo 1) .....	20
Estimadores de dispersão residual para incremento de uniformidade considerando efeito aditivo de pai de touro (Modelo 2).....	21
Análise de resíduos em escala categórica (Modelo 3) .....	23
2.3. RESULTADOS E DISCUSSÃO .....	25
Primeiro passo das análises .....	25
Segundo passo das análises .....	25
2.4. CONCLUSÕES.....	33
2.5. REFERÊNCIAS .....	33
CAPÍTULO 3 - ABORDAGEM EM DOIS PASSOS PARA ESTUDAR A VARIABILIDADE GENÉTICA PARA DISPERSÃO RESIDUAL NO GANHO DE PESO DO NASCIMENTO À DESMAMA NA RAÇA NELORE .....	36
3.1. INTRODUÇÃO.....	37
3.2. MATERIAL E MÉTODOS .....	38
3.3. RESULTADOS e DISCUSSÃO .....	43
Estatísticas descritivas referentes às observações de GND e de sua dispersão residual .....	43
Parâmetros genéticos para a média de GNDa .....	45
Evidências de efeitos aditivos sobre a dispersão residual de GNDa.....	46
Estudo sobre possíveis efeitos de prepotência dos touros sobre a variabilidade da progênie.....	54
3.4. CONCLUSÕES.....	56
3.5. REFERÊNCIAS .....	58
4.1. INTRODUÇÃO.....	62
4.2. MATERIAL E MÉTODOS .....	62
Estrutura da população simulada .....	62
Modelos simulados .....	63
Cenários simulados e critérios de comparação .....	64

Estimação de componentes de variância para a média e variância residual da característica simulada .....	65
Modelo alternativo para o segundo passo das análises .....	66
4.3. RESULTADOS E DISCUSSÃO .....	67
4.4. CONCLUSÕES .....	77
4.5. REFERÊNCIAS .....	78
CAPÍTULO 5. COMPARAÇÃO DE MODELOS ALTERNATIVOS CONSIDERANDO HETEROGENEIDADE GENÉTICA E AMBIENTAL SOBRE A VARIÂNCIA RESIDUAL DO GANHO DE PESO DO NASCIMENTO À DESMAMA DE BOVINOS DA RAÇA NELORE .....	
NELORE .....	79
5.1. INTRODUÇÃO .....	80
5.2. MATERIAL E MÉTODOS .....	80
5.3. RESULTADOS E DISCUSSÃO .....	84
5.4. CONCLUSÕES .....	93
5.5. REFERÊNCIAS .....	93
CAPÍTULO 6. IMPLICAÇÕES .....	95
APÊNDICE I – DERIVAÇÃO DO FATOR $f$ , USADO PARA TRANSFORMAR AS VARIÂNCIAS ADITIVAS DOS ESTIMADORES DE ESCALA PARA A ESCALA DE VARIÂNCIA .....	96

## CRITÉRIOS DE SELEÇÃO PARA INCREMENTO DE UNIFORMIDADE DE PRODUÇÃO EM BOVINOS DE CORTE

**RESUMO** – O objetivo deste estudo foi investigar a existência de variabilidade genética aditiva sobre a variância residual do ganho de peso do nascimento à desmama (GND) de bovinos Nelore e as perspectivas de se explorar diferenças entre genótipos para variância residual para a obtenção de maior uniformidade de produção, por meio de seleção. Diferentes abordagens, implementadas em dois passos, foram estudadas: Inicialmente, avaliaram-se três modelos para análise de medidas de dispersão dos resíduos associados às observações de GND da progênie de touros Nelore. O modelo considerado mais promissor foi empregado em estudo subsequente, em que foi investigado o impacto do tamanho de progênie dos touros nas estimativas obtidas para variância aditiva sobre a dispersão residual e estimadores de dispersão em diferentes escalas foram comparados. A confiabilidade de tal abordagem foi verificada por meio de simulação de Monte Carlo. Um último estudo avaliou a possibilidade de se considerarem, simultaneamente, efeitos aditivos e ambientais sobre a variância residual de GND, empregando-se diferentes modelos para análise do logaritmo natural do quadrado do resíduo associado a cada observação. Concluiu-se que, ao se considerar famílias de grande tamanho, seria possível obter predições acuradas do mérito genético dos touros para a variância residual e alguma resposta em termos de uniformidade de produção, sendo a abordagem do último estudo considerada a mais adequada para este fim. Desconsiderar efeitos ambientais sobre a variância residual no segundo passo das análises pode levar a superestimação da variância aditiva sobre a dispersão residual, bem como da resposta esperada à seleção.

**Palavras-chave:** estimadores de dispersão, heterocedasticidade, valor genético, variância residual

## **SELECTION CRITERIA TO IMPROVE UNIFORMITY OF BEEF CATTLE PRODUCTION**

**ABSTRACT** – This study was carried out to investigate the existence of genetic variability on residual variance of beef cattle production traits and to evaluate the opportunity for improvement in uniformity of such traits by selecting for lower residual variance. Different two-step approaches were studied to address these questions: Firstly, three models were employed to analyze different measures associated with residual dispersion of weight gain from birth to weaning (GND) in the progeny of Nellore sires. The model that performed best was employed in a subsequent study to assess the impact of progeny size on estimates of additive variance for residual dispersion, also aiming to compare dispersion estimators of different scales and to predict selection response in each situation. Reliability of this approach was verified by Monte Carlo simulation. The possibility of considering, simultaneously, additive and environmental effects on residual variance of GND was investigated by analyzing log squared residuals associated with each observation according to different models. It was concluded that, by considering large sire families, accurate estimates of genetic merit of sires for residual variance could be obtained as well as some improvement in uniformity of GND. Analyzing log squared residuals associated with each observation was considered the most promising approach for this task. Ignoring environmental effects at the level of residual variance could lead to inflated estimates of additive variance of residual dispersion, therefore implying in overestimation of response to selection.

**Keywords:** breeding value, dispersion estimators, heteroscedasticity, residual variance

## **CAPÍTULO 1. CONSIDERAÇÕES GERAIS**

O tema estudado nesta dissertação envolve o emprego de diferentes abordagens para estimação de parâmetros genéticos sobre a variância residual de características de produção de bovinos de corte. O texto está dividido em seis capítulos, brevemente descritos a seguir.

O presente capítulo compreende uma introdução ao tema, seguida de uma revisão de literatura e da apresentação dos objetivos do trabalho.

O capítulo 2 sumariza os resultados dos estudos preliminares deste trabalho, visando comparar diferentes modelos para estimação de variância aditiva sobre a dispersão residual.

No capítulo 3, a abordagem identificada como mais promissora no estudo do capítulo 2 foi investigada de modo mais aprofundado, verificando-se a influência de diferentes fatores sobre a variância aditiva estimada para dispersão residual e algumas perspectivas de resposta à seleção. No capítulo 4, apresentam-se os resultados de um estudo de simulação desenvolvido para validação da abordagem empregada no capítulo 3.

No capítulo 5, a possibilidade de emprego de um modelo alternativo que permitiria contemplar simultaneamente efeitos ambientais e aditivos sobre a variância residual foi investigada, comparando-se os resultados obtidos com aqueles dos capítulos anteriores.

No capítulo 6, apresentam-se as considerações finais sobre o trabalho, incluindo algumas implicações, limitações identificadas com o uso das abordagens propostas e sugestões de futuros estudos.

### *1.1. Introdução*

Tem sido apontada tendência de crescimento na demanda mundial por carne bovina, bem como por outras fontes de proteína animal, em virtude do crescimento e da melhora do poder aquisitivo médio da população mundial. O Brasil é atualmente o maior exportador mundial de carne bovina. Porém, para consolidar esta posição e, sobretudo, conquistar novos mercados, os produtores têm que se adaptar a crescentes exigências em termos de

qualidade de produto, sustentabilidade ambiental, responsabilidade social e competitividade de preços.

No aspecto de qualidade de produto final, sobretudo com vistas à uniformidade e consistência no atendimento aos padrões determinados pelos diferentes mercados, percebe-se que tais exigências têm provocado readequações nos sistemas de produção de alimentos, inclusive no âmbito da produção animal (GIOVANNUCCI & REARDON, 2000; TRIENEKENS & ZUURBIER, 2008).

Como exemplos de tal tendência na pecuária, podem-se citar políticas de bonificação/penalização sobre o preço da arroba de carcaça em função do atendimento a padrões ótimos (KANIS et al., 2006), ou mesmo a oportunidade de remuneração diferenciada para a carne certificada segundo os padrões estabelecidos por alguns grupos privados (e.g. *Certified Angus Beef*, CAB, 2009).

Nas últimas décadas, as metodologias de predição de valores genéticos e seleção têm sido eficientes em incrementar o desempenho produtivo médio dos rebanhos. Porém, segundo GARRICK & ENNS (2003), os programas de melhoramento de bovinos de corte do mundo inteiro têm praticado seleção com considerável ênfase no peso vivo, o que tem levado a incrementos nas taxas de crescimento e nos pesos em quase todas as idades, incluindo peso ao nascer e peso adulto de vacas, o que seria indesejável, dado o impacto negativo no desempenho reprodutivo e custo de manutenção de matrizes, respectivamente.

Deste modo, em muitas situações, há necessidade de aumento da proporção de animais com desempenho dentro de um intervalo ótimo em termos econômicos e/ou funcionais (aumento da uniformidade) e não apenas o incremento do desempenho médio dos animais.

HOHENBOKEN (1985) reviu o potencial de algumas estratégias de acasalamento e de seleção para diminuir a variabilidade fenotípica em animais domésticos. Todavia, apenas o uso de seleção possibilitaria mudança permanente sobre a uniformidade de produção dos animais.

Dentre algumas alternativas que poderiam permitir redução na variabilidade fenotípica por meio de seleção, pode-se citar a consideração de diferenças genéticas entre indivíduos para a variância residual nas avaliações genéticas (e.g. SORENSEN & WAAGENPETERSEN, 2003; MULDER et al. 2007) e o uso de estimadores alternativos que

permitissem discriminar diferenças na variabilidade da progênie dos animais candidatos (ORTIZ-PEÑA et al., 2004; MUNIZ et al., 2005; FRIES et al., 2006; ZAMPAR, 2007).

### *1.2. Modelos considerando heterogeneidade genética sobre a variância residual*

Segundo SORENSEN & WAAGENPETERSEN (2003), a hipótese de parte da variância residual ser controlada geneticamente abre duas importantes questões no âmbito do melhoramento animal. A primeira é que a predição de resposta à seleção segundo o modelo clássico pode não ser tão acurada, sendo necessário investigar a relevância desse possível erro. A outra é a possibilidade de avaliar o efeito da seleção para uma característica sobre sua variabilidade fenotípica.

O uso de metodologias que explorassem a variabilidade genética da variância residual ( $\sigma^2_E$ ) poderia permitir a obtenção de resposta sobre a variabilidade fenotípica, sendo de grande valia para incrementar a uniformidade ou mesmo para impor restrições sobre a variabilidade.

Ao contrário do pressuposto de igualdade de variância residual entre genótipos, admitido em avaliações tradicionais, diferentes modelos que assumem heterogeneidade genética sobre a variância residual têm sido propostos. Basicamente, tais modelos podem ser classificados em dois tipos:

Modelos aditivos: incluem efeitos genéticos aditivos sobre a média e a variância residual (e.g. HILL & ZHANG, 2004; MULDER et al., 2007);

Modelos exponenciais: incluem efeitos aditivos sobre a média e o logaritmo natural da variância residual (SANCRISTOBAL-GAUDY et al., 1998; SORENSEN & WAAGEPETERSEN, 2003).

A estrutura básica desses modelos pode ser representada por:

$$P = \mu + A_m + \chi * \text{sqrt}(\sigma^2_E + A_v) \quad (\text{Modelo aditivo})$$

$$P = \mu + A_m + \chi * \exp \{[\ln(\sigma^2_E) + A_v] / 2\} \quad (\text{Modelo exponencial}), \text{ em que:}$$

P é o fenótipo da característica,  $\mu$  é a média da característica;  $A_m$  e  $A_v$  são os valores genéticos individuais, respectivamente, para a média e para a variância residual;  $\sigma^2_E$  é a variância residual quando  $A_v = 0$  e  $\chi$  é uma variável com distribuição normal padrão,  $\chi \sim N(0,1)$ .

MULDER et al. (2007) apontaram similaridades entre estes dois tipos de modelos e propuseram uma equação para converter as variâncias genéticas obtidas usando modelo exponencial em um modelo aditivo, uma vez que este último apresentaria melhor adequação com a teoria genética quantitativa. Sob a pressuposição de aditividade, estes autores sugeriram uma medida de herdabilidade da variância residual ( $h^2_v$ ) calculada como a inclinação da regressão do quadrado do fenótipo sobre  $A_v$ :

$$h^2_v = \sigma^2_{A_v} / (2 * \sigma^4_P + 3 * \sigma^2_{A_v}), \text{ em que:}$$

$\sigma^2_{A_v}$  é a variância genética sobre a variância residual segundo o modelo linear e  $\sigma^4_P$  é o quadrado da variância fenotípica da característica.

Segundo MULDER et al. (2007), o parâmetro  $h^2_v$  estaria associado à acurácia de predição do mérito genético para variância residual. Estes mesmos autores sugeriram como medida da variabilidade genética para variância residual, o parâmetro *evolvability* ( $Ev$ ) (HOULE, 1992):

$$Ev = \sigma_{A_v} / \mu,$$

em que  $\sigma_{A_v}$  é o desvio padrão aditivo para variância residual e  $\mu$  é a dispersão residual média, de modo que  $Ev$  seria um parâmetro que mediria a oportunidade de resposta à seleção.

Sob o modelo aditivo, seria possível calcular a resposta predita da seleção para menor  $A_v$  considerando os valores de acurácia da seleção ( $r_{\hat{A}_v, A_v}$ ) calculada com base em  $h^2_v$ , a *evolvability* ( $Ev$ ) da variância residual de uma dada característica e a intensidade de seleção ( $i$ ), permitindo prever a redução relativa esperada na variância residual média ( $\Delta\sigma^2_E / \sigma^2_E$ ), após uma geração de seleção, como uma extensão da equação clássica de predição de resposta à seleção (BIJMA, 2009). Assim:

$$\Delta\sigma^2_E / \sigma^2_E = r_{\hat{A}_v, A_v} * i * Ev.$$

Um dos primeiros modelos quantitativos a admitir heterogeneidade para variância ambiental e outros componentes de variância foi proposto por FOULLEY & QUAAS (1995). SANCRISTOBAL-GAUDY et al. (1998) incorporaram a este modelo a influência de efeitos genéticos aditivos sobre a variância ambiental, também implementada por meio de metodologia *bayesiana* no estudo de SORENSEN & WAAGEPETERSEN (2003).

IBÁÑEZ-ESCRICHE et al. (2008) usaram um modelo similar ao descrito por SANCRISTOBAL-GAUDY et al. (1998) para analisar dados de peso corporal de suínos aos

175 dias de idade, além de terem investigado conseqüências econômicas da existência de heterogeneidade de variância residual geneticamente estruturada.

Outra proposta de modelo considerando heterogeneidade genética sobre a variância residual é analisar, para cada observação de uma determinada característica, o logaritmo natural do quadrado dos resíduos preditos,  $\ln(\hat{\epsilon}^2)$ , assumindo esta medida como um estimador da variância residual. Este tipo de modelo tem sido implementado por meio de uma abordagem em dois passos e reduziria, consideravelmente, a complexidade dos algoritmos de estimação comparada a de modelos implementados num único passo para o mesmo fim (e.g. GARREAU et al., 2008; MULDER et al., 2009).

### 1.3. Evidências de heterogeneidade genética sobre a variância residual

O estudo conduzido por Mackay & Lyman (2005) tem sido citado por diversos trabalhos como uma das evidências mais contundentes da existência de heterogeneidade genética sobre a variância residual. Estes autores encontraram considerável variabilidade entre linhas isofêmeas (clones) de *Drosophila melanogaster* para a variância residual dentro de linha, considerando as características número de cerdas abdominais e número de cerdas esternopleurais, o que forneceu um importante suporte do ponto de vista experimental para este fenômeno biológico.

No caso de animais de produção, em vários estudos com diferentes espécies tem sido relatada a existência de heterogeneidade genética na variância residual ( $\sigma^2_E$ ), como no caso de peso corporal de frangos (ROWE et al., 2006), de prolificidade em ovinos (SANCRISTOBAL-GAUDY et al., 2001) e em suínos (SORENSEN & WAAGENPETERSEN, 2003) e de produção de leite em bovinos (VAN VLECK, 1968).

ROWE et al. (2006) utilizaram famílias de grande tamanho e condições ambientais uniformes num experimento com frangos de corte para estimar a variância dentro de cada família de machos, em dados relativos ao peso dos animais aos 35 dias de idade. Os resultados obtidos permitiram identificar a presença de heterogeneidade genética sobre  $\sigma^2_E$ .

Em suínos, têm-se encontrado valores modestos, porém significativos, de herdabilidade para a variabilidade dentro de famílias em características ligadas ao peso corporal (DAMGAARD et al., 2003) e composição de carcaça (ZETTELER & KNOL, 2005), sugerindo que alguns machos podem apresentar progênie com desempenho mais uniforme.

KANIS et al. (2006), ao contrário, não encontraram evidências de que a variabilidade dentro de famílias para peso vivo ao abate e peso da carcaça quente em suínos seja herdável.

Segundo SANCRISTOBAL-GAUDY et al. (2001), a seleção para incremento da prolificidade em ovinos aumenta, concomitantemente, sua variabilidade, ao passo que o número de dois cordeiros por parto é considerado ótimo. Para esta característica, estes autores encontraram expressiva heterogeneidade de variância dentro da progênie de reprodutores, o que poderia permitir canalizar a seleção para a prolificidade em torno do ótimo.

Os resultados do estudo de GARREAU et al. (2008), após quatro gerações de seleção divergente para variabilidade do peso ao nascer (PN) de coelhos, evidenciaram resposta significativa sobre o desvio-padrão do peso ao nascer dentro de ninhada, observando-se um valor 14,8% menor na linhagem homogênea, que havia sido selecionada para menor variância residual em PN, usando-se a medida  $\ln(\hat{\epsilon}^2)$  como estimador deste parâmetro. Além disso, estes autores observaram que a linhagem mais homogênea apresentou mortalidade significativamente menor no período até a desmama, bem como maior número médio de animais desmamados (1,83 coelhos a mais).

Além da existência de variabilidade genética para  $\sigma^2_E$ , outra questão importante a ser estudada é a associação entre efeitos genéticos sobre a média e a variância residual das características de produção.

Na literatura, não há evidências de um padrão geral para este tipo de parâmetro. Em alguns estudos, estimaram-se correlações elevadas e positivas entre estes efeitos (ROS et al., 2004; GUTIERREZ et al. 2006), ao passo que estimativas negativas de magnitude igualmente elevada também foram relatadas por outros autores (e.g. HUBY et al., 2003; GUTIERREZ et al. 2006). Apesar da evidência de heterogeneidade genética de  $\sigma^2_E$  encontrada no estudo de IBÁÑEZ-ESCRICHE et al. (2008), estes autores não encontraram correlação significativa entre o efeito aditivo sobre a média ( $A_m$ ) e o efeito aditivo sobre a variância residual ( $A_v$ ).

Tais diferenças nas estimativas de correlação relatadas em diferentes trabalhos, especialmente no caso de características similares, indicariam a necessidade de se investigar melhor a influência de efeitos genéticos, ambientais, bem como de efeitos de escala sobre este parâmetro.

No que tange a efeitos de escala, YANG et al. (2009) alertaram para o fato de que, como a assimetria da distribuição marginal dos dados no modelo heteroscedástico é diretamente proporcional ao coeficiente de correlação entre  $A_m$  e  $A_v$ , as estimativas de parâmetros genéticos sob este tipo de modelo poderiam ser influenciadas por efeitos de escala.

Segundo estes autores, após uso de transformação Box-Cox em dados de tamanho de leitegada em coelhos (visando remoção de efeitos escala), as estimativas obtidas para  $\sigma^2_{Av}$  e da correlação entre  $A_m$  e  $A_v$  foram alteradas significativamente, obtendo-se evidências bem menos contundentes de heterogeneidade genética sobre  $\sigma^2_E$ , comparando com os resultados correspondentes obtidos antes da transformação.

Na Tabela 1 (adaptada de MULDER et al., 2007), apresentam-se estimativas ligadas à variância genética sobre a variância residual encontradas em cinco estudos. Estes autores usaram tais resultados para padronizar as informações provenientes de diferentes modelos, calculando a variância genética sobre a variância residual segundo modelo aditivo ( $\sigma^2_{Av}$ ).

Há evidências de que a magnitude da herdabilidade da variância residual seja pequena, indicando necessidade de grande quantidade de informação para predizer de modo acurado o valor genético para  $\sigma^2_E$ . Por outro lado, os valores elevados de  $E_v$  (*evolvability*) indicariam possibilidade de obtenção de resposta razoável sobre a uniformidade de produção por meio de seleção (Tabela 1).

MULDER et al. (2008) estudaram a influência de alguns parâmetros genéticos e número de progênies por reprodutor na resposta da média e da variância do peso da carcaça de suínos, admitindo heterogeneidade genética sobre  $\sigma^2_E$  e aplicando um índice (combinando  $A_m$  e  $A_v$ ) para selecionar os animais. A resposta predita para  $\sigma^2_E$  após uma geração, com no mínimo 50 progênies por macho, foi considerável (redução de mais de 10% da variância residual na geração anterior) e o esquema de teste de progênie foi considerado o mais eficiente para se reduzir  $\sigma^2_E$  por meio de seleção. Os autores ressaltaram que seleção visando uniformidade seria interessante principalmente para características cuja média já se encontrasse próxima do “ótimo”.

Tabela 1 . Comparação de diferentes estimativas de variância genética sobre a variância residual

Fonte	Característica	$\sigma_{Av, exp}^2$	$\sigma_{Av}^2$	$h^2_v$	Ev
SAN-CRISTOBAL GAUDY et al. (1998)	pH da carne suína	0,150	1,23e <sup>-4</sup>	0,039	0,402
SAN-CRISTOBAL GAUDY et al. (2001)	prolificidade em ovinos	0,230	0,057	0,048	0,509
SORENSEN & WAAGENPETERSEN (2003)	prolificidade em suínos	0,090	4,921	0,026	0,307
ROS et al. (2004)	peso corporal em <i>escargot</i>	0,290	0,368	0,017	0,580
ROWE et al. (2006)	peso corporal em frangos (machos)	0,086	8460	0,029	0,299
	peso corporal em frangos (fêmeas)	0,096	5310	0,031	0,318

$\sigma_{Av, exp}^2$  é a variância genética sobre a variância residual segundo modelo exponencial (usado originalmente nestes estudos).  $Ev = \sigma_{Av} / \sigma_{E,}^2$ , denominado *evolubility*, medida proposta por Houle (1992) para medir a oportunidade de resposta à seleção.

#### 1.4. Medidas de dispersão e uso de estimadores alternativos às DEPs para seleção contemplando maior uniformidade

FRIES et al. (2006) levantaram a questão de que o uso atual de estimadores de tendência central (DEPs ou EBVs, obtidos por equações de modelos mistos) não seria eficiente para implementar seleção contemplando o atendimento consistente a padrões de qualidade estabelecidos, uma vez que tais estimativas não consideram as dispersões dentro de subclasses. O uso de estimadores que consideram a dispersão na progênie dos touros tem sido investigado em alguns estudos visando contornar este problema (e.g. MUNIZ et al., 2005; FRIES et al., 2006; TANAKA et al., 2009).

Dentre as medidas de dispersão utilizadas em programas de melhoramento genético, destaca-se o desvio-padrão (SD), medida considerada estável e expressa na mesma escala da característica original. Contudo, é sensível a valores discrepantes, uma vez que elevar ao quadrado os desvios das observações, aumenta exponencialmente cada unidade de distância em relação à média, efeito que não é totalmente removido pelo cálculo da raiz quadrada, como discutido por ROBERSON et al. (2007).

Um conceito importante associado à robustez de uma medida é o *breakdown point* (BP), definido por GNANADESIKAN (1997) como a maior fração das observações em uma amostra em que pode haver valores extremos sem distorcer o valor do estimador. Uma alternativa robusta para medida de dispersão, principalmente em presença de valores extremos no conjunto de observações, é a mediana do desvio absoluto da mediana (*median absolute deviation*, MAD), proposta por HAMPEL (1974) e calculada por:

$$\text{MAD} = \text{Mediana} (|X_i - \text{Mediana}(X)|), \text{ em que:}$$

$X_i$  é o valor da  $i$ -ésima observação para a variável e  $\text{Mediana}(X)$  é a mediana da variável.

Apesar de a MAD ser um estimador de escala muito robusto (BP=50%), ROUSSEEUW & CROUX (1993) apontaram sua baixa eficiência (calculada com base na variância assintótica do estimador) para dados com distribuições de fato normais (37%) e o pressuposto implícito de simetria como suas limitações. Os autores propuseram duas medidas robustas de dispersão (*Sn scale* e *Qn scale*), alternativas à MAD.

A medida  $S_n$  é calculada por:

$$S_n = c * \text{Mediana}_i \{ \text{Mediana}_j |x_i - x_j| \}, \text{ em que:}$$

para a  $i$ -ésima observação computa-se a mediana de  $\{|x_i - x_j| ; j = 1, \dots, n\}$  e  $S_n$  é a mediana das  $n$  medianas calculadas, sendo  $c$  uma constante usada para fazer de  $S_n$  um estimador consistente (para dados normais,  $c=1,1926$ ).

$Q_n$  pode ser calculada como o quantil 0,25 das distâncias:

$$\{|x_i - x_j| ; i \leq j\},$$

e multiplicada por  $d$ , constante com função semelhante à da constante  $c$  de  $S_n$  e cujo valor é 2,2219, para dados normais.

ROUSSEEUW & CROUX (1993) verificaram que a *Sn scale* possui as mesmas propriedades de robustez da MAD quanto ao BP, mas apresenta maior eficiência em distribuições normais (58%) e não depende de simetria, uma vez que  $S_n$  mede distâncias entre valores (ao contrário de MAD e SD, que medem distância em relação a uma medida de tendência central). A outra medida proposta por estes autores, chamada de *Qn scale*, se diferencia de  $S_n$  pela sua maior eficiência para dados normais (86%) e maior tempo necessário para computar seu algoritmo.

A reexpressão da característica ganho de peso médio diário (GMD) como uma função de sua média harmônica seria um exemplo de uso de estimadores alternativos para

incremento de uniformidade de produção, geralmente definido como número de dias para atingir um peso alvo (DIAS). Tal abordagem tem sido apontada como eficiente para selecionar touros com progênie mais uniforme (ORTIZ-PEÑA et al., 2004; MUNIZ et al., 2005; TANAKA et al., 2009). Porém, SIROL et al. (2007) verificaram por meio de simulação que observações de GMD com valores muito baixos em um grupo de contemporâneos (GC) poderiam ter influência importante na avaliação com base em DIAS dos touros com filhos naquele grupo, dependendo da estrutura dos GC.

CARVALHEIRO et al. (2001), em estudo de simulação considerando ganho de peso do nascimento à desmama de bovinos, demonstraram que entre 95% e 99% das diferenças entre as estimativas para GMD e DIAS podiam ser explicadas pelas diferenças na variabilidade da progênie de touros.

ZAMPAR (2007) verificou que o uso da média harmônica para estimar DEPs a fim de ajustar o valor genético predito de touros para a dispersão dos valores genéticos de suas progênies não implicou em diferenças significativas, em comparação à estimativa por média aritmética e enfatizou a necessidade de mais estudos a fim de se encontrar uma metodologia que permita incrementar a uniformidade da progênie de touros. Cabe ressaltar que em tal estudo utilizou-se um procedimento diferente daqueles adotados por CARVALHEIRO et al. (2001), ORTIZ-PEÑA et al. (2004) e MUNIZ et al. (2005) para obter DEPs como uma função da média harmônica.

ZAMPAR (2007) estudou as diferenças de classificação dos touros com base em critérios considerando as médias aritmética e harmônica das DEPs dos filhos de cada touro, obtidas por meio de um modelo animal, considerando as características peso à desmama, peso ao sobreano e ganho de peso pós-desmama. Todavia, em estudos como o de ORTIZ-PEÑA et al. (2004), parâmetros genéticos para a característica DIAS foram estimados diretamente, sendo as DEPs obtidas sob tais modelos usadas para comparar a classificação dos touros com aquela obtida com a DEP convencional. Além disso, a variabilidade da progênie dos touros foi calculada com base nas próprias DEPs por ZAMPAR (2007) e, com base em resíduos (observações ajustadas para efeitos ambientais e maternos) por ORTIZ-PEÑA et al. (2004).

FRIES et al. (2006) investigaram o uso de dois estimadores alternativos para discriminar variabilidade da progênie de touros e incrementar a uniformidade por meio de

seleção. Foram consideradas 14 características, sendo um dos estimadores baseado no uso da média harmônica e o outro no uso de um parâmetro de dispersão similar à MAD (*average absolute deviation*, AAD), definido por:  $AAD = \text{Média } |X_i - m(X)|$ , em que  $X_i$  é o valor da  $i$ -ésima observação para a variável e  $m(X)$  é a medida de tendência central da variável. Neste estudo, o estimador baseado em AAD manteve propriedades desejáveis para as características estudadas, ao passo que o primeiro foi considerado sensível a valores próximos de zero dos resíduos usados para seu cálculo.

Conforme discutido em diferentes estudos, um dos principais problemas nas análises de heterogeneidade genética sobre a variância residual é sua estimabilidade (e.g. MULDER et al., 2009; WOLC et al., 2009). Talvez por esta razão diferentes modelos e procedimentos tenham sido investigados em estudos desta natureza, muitos dos quais têm sido desenvolvidos como meio de contornar limitações computacionais relatadas em alguns trabalhos.

Como parece não haver consenso na literatura quanto ao modelo e procedimento mais adequado para investigar heterogeneidade genética sobre  $\sigma^2_E$  e dado o crescente interesse por maior uniformidade na produção animal, fica evidente a necessidade de estudos (incluindo experimentos de seleção) que permitam avanços no conhecimento sobre este tipo de questão, sendo recomendável, inclusive, que diferentes métodos sejam comparados, de modo a identificarem-se abordagens que mais se adequem a cada estrutura populacional e tipo de característica.

### 1.5. Objetivos

O presente trabalho foi conduzido com os objetivos de: i) investigar a existência de variabilidade genética aditiva sobre a variância residual do ganho de peso do nascimento à desmama em bovinos de corte; ii) estudar uma eventual associação de ordem genética entre efeitos aditivos sobre a média e a variância residual desta característica; iii) investigar algumas perspectivas de seleção contemplando maior uniformidade de produção, considerando diferentes cenários de importância econômica para este critério.

### 1.6. Referências

- BIJMA, P. Quantitative genetic models for social interaction and GxE and inherited variability. Quantitative Genetic Theory and Analysis (Course Notes), Armidale Animal Breeding Summer Course, University of New England, Armidale, 2009.
- CAB. 2009. Certified Angus Beef Homepage. Disponível em: <http://www.certifiedangusbeef.com>. Acesso em: 12.nov.2009.
- CARVALHEIRO, R.; MUNIZ, C. A. S. D.; QUEIROZ, S. A. et al. Estudo de simulação das relações entre a média aritmética, a média harmônica e o desvio-padrão. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 38., 2001, Piracicaba. Anais... Piracicaba: SBZ, 2001., p.627.
- DAMGAARD, L.H.; RYDHMER, L.; LØVENTAL, P., GRANDINSON, K. Genetic parameters for within-litter variation in piglet birth weight and change in within-litter variation during suckling. **Journal of Animal Science**, Savoy, v. 81, pp. 604-610, 2003.
- FOULLEY, F. L.; QUAAS, R. L. Heterogeneous variances in Gaussian linear mixed models. **Genetics Selection Evolution**, n. 27, pp. 211–228, 1995.
- FRIES, L.A.; BRITO, F.V.; CARDOSO, V. et al. Preliminary assessment of robust joint location & dispersion estimators of breeding values as applied to beef cattle traits in Brazil. In: 8<sup>TH</sup> WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 2006, Belo Horizonte. **Proceedings...**Belo Horizonte:WCGALP, 2006.
- GARREAU, H.; BOLET, G.; LARZUL, C.; ROBERT-GRANIÉ, C.; SALEIL, G.; SANCRISTÓBAL-GAUDY, M.; BODIN, L. Results of four generations of a canalising selection for rabbit birth weight. **Livestock Science**, v. 119, pp. 55-62, 2008.
- GARRICK, D.J.; ENNS, R.M. How best to achieve genetic change. In: Beef Imp. Fed. 35th Annu. Res. Symp. & Annu. Mtg., 2003, Lexington. **Proceedings...**Lexington: BIF, 2003, pp. 28–31.
- GIOVANNUCCI, D.; REARDON, T. **Understanding Grades and Standards and How to Apply Them: The Guide to Developing Agricultural Markets and Agro-enterprises**, 2000. Disponível em: [http://marketstandards.chemonics.net/resources/Critical%20Reports/grades\\_and\\_stds.pdf](http://marketstandards.chemonics.net/resources/Critical%20Reports/grades_and_stds.pdf). Acesso em: 08.fev.2008.

- GNANADESIKAN, R. **Methods for statistical data analysis of multivariate observations**. New Jersey: John Wiley, 1997. 353 p.
- GUTIERREZ, J. P.; NIETO, B.; PIQUERAS, P.; IBAÑEZ, N.; SALGADO, C. Genetic parameters for canalisation analysis of litter size and litter weight traits at birth in mice. **Genetics Selection Evolution**, Les Ulis, v. 38, pp. 445–462, 2006.
- HAMPEL, F. The influence curve and its role in robust estimation. **Journal of the American Statistical Association**, Alexandria, v 69, pp. 383–393, 1974.
- HILL, W. G.; ZHANG, X. S. Effects of phenotypic variability of directional selection arising through genetic differences in residual variability. **Genetics Research**, Cambridge, n.83, pp. 121–132, 2004.
- HOHENBOKEN, W. D. The Manipulation of Variation in Quantitative Traits: A Review of Possible Genetic Strategies. **Journal of Animal Science**, Savoy, n. 60, pp. 101-110, 1985.
- HOULE, D. Comparing evolvability and variability of quantitative traits. **Genetics**, Bethesda, v. 130, pp. 195-204, 1992.
- HUBY M.; GOGUÉ, J.; MAIGNEL, L.; BIDANEL, J.P. Corrélations génétiques entre les caractéristiques numériques et pondérales de la portée, la variabilité du poids des porcelets et leur survie entre la naissance et le sevrage. **Journées Recherche Porcine**, v. 35, pp. 293–300, 2003.
- IBAÑEZ-ESCRICHE, N.; VARONA, L., SORENSEN, D.; NOGUERA, J.L. A study of heterogeneity of environmental variance for slaughter weight in pigs. **Animal**, Cambridge, n.2, v.1, pp.19-26, 2008.
- KANIS, E.; VAN PELT, M. L; BONEKAMP, P. R. T.; KNOL, E. F. Is within-family variation in carcass weight of pigs heritable? In: 8<sup>TH</sup> WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 2006, Belo Horizonte. **Proceedings...Belo Horizonte:WCGALP**, 2006.
- MACKAY, T. F. C.; LYMAN, R. F. Drosophila bristles and the nature of quantitative genetics variation. **Philosophical Transactions of the Royal Society of London, Series B**, v. 360, pp.1513–1527, 2005.
- MULDER, H. A., HILL, W. G.; VEREIJKEN, A.; VEERKAMP, R. F. Estimation of genetic variation in residual variance in female and male broiler chickens. **Animal**, v. 3, pp. 1673-1680, 2009.

- MULDER, H. A.; BIJMA, P.; HILL, W. G. Prediction of Breeding Values and Selection Responses With Genetic Heterogeneity of Environmental Variance. **Genetics**, Bethesda, v.175, pp. 1895-1910, 2007.
- MULDER, H. A.; BIJMA, P.; HILL, W. G. Selection for uniformity in livestock by exploiting genetic heterogeneity of residual variance. **Genetics Selection Evolution**, Les Ulis, v. 40, n. 1, pp. 37-59, 2008.
- MUNIZ, C.A.S.D.; CARVALHEIRO, R.; FRIES, L.A.; QUEIROZ, S.A. Dois critérios de seleção na pré-desmama em bovinos da raça Gir. 2. Efeito na classificação dos animais. **Revista Brasileira de Zootecnia**, Viçosa, v.34, n.3, p.816-826, 2005.
- ORTIZ PEÑA, C.D.; CARVALHEIRO, R.; QUEIROZ, S.A.; FRIES, L.A. Comparison of selection criteria for pre-weaning growth traits of Nelore cattle. **Livestock Production Science**, Amsterdam, v.86, p.163-167, 2004.
- ROBERSON, Q. M.; STURMAN, M. C.; SIMONS, T. L. Does the Measure of Dispersion Matter in Multilevel Research? A Comparison the Relative Performance of Dispersion Indexes. **Organizational Research Methods**, v.4, n. 10, 2007.
- ROS, M.; SORENSEN, D.; WAAGEPETERSEN, R.; DUPONT-NIVET, M.; BONNET, J. C.; MALLARD, J. Evidence for Genetic Control of Adult Weight Plasticity in the Snail *Helix aspersa*. **Genetics**, Bethesda, v. 168, pp. 2089-2097, 2004.
- ROUSSEEUW, P. J.; CROUX C. Alternatives to the Median Absolute Deviation. **Journal of the American Statistical Association**, Alexandria, v. 88, n. 424, p. 1273-1283, 1993.
- ROWE, S. J.; WHITE, I. M.S., AVENDAÑO, S.; HILL, W. G. Genetic heterogeneity of residual variance in broiler chickens. **Genetics Selection Evolution**, Les Ulis, n. 38, pp. 617-635, 2006.
- SANCRISTOBAL-GAUDY, M.; BODIN, L.; ELSESEN, J-M.; CHEVALET, C. Genetic components of litter size variability in sheep. **Genetics Selection Evolution**, Les Ulis, n. 33, pp.249–271, 2001.
- SANCRISTOBAL-GAUDY, M.; ELSESEN, J-M.; BODIN, L.; CHEVALET, C. Prediction of the response to a selection for canalisation of a continuous trait in animal breeding. **Genetics Selection Evolution**, Les Ulis, n. 30, pp. 423–451, 1998.

- SIROL, M.L.F.G, HADLICH, J.C.; SUGUISAWA, L.; ABRAHÃO, A.R.; OLIVEIRA, H.N. Efeito de valores discrepantes em grupos de contemporâneos na predição dos valores genéticos. **Revista Brasileira de Zootecnia**, Viçosa, v.36, n.4, p.807-812, 2007.
- SORENSEN, D.; WAAGEPETERSEN, R. Normal linear models with genetically structured residual variance heterogeneity: a case study. **Genetics Research**, Cambridge, v. 82, pp. 207–222, 2003.
- TANAKA, A. L. R.; CARVALHEIRO, R.; FRIES, L.A.; QUEIROZ, S.A. Comparação de critérios de seleção para precocidade de crescimento em bovinos da raça Guzerá. **Revista Brasileira de Zootecnia**, v. 39, n.2., pp. 284-291, 2009.
- TRIENEKENS, J.; ZUURBIER, P. Quality and safety standards in the food industry, developments and challenges. **International Journal of Production Economics**, Amsterdam, v. 113, n .1, pp. 107-122, 2008.
- VAN VLECK, L. D. Variation of Milk Records Within Paternal-Sib Groups. **Journal of Dairy Science**, Savoy, v. 51 n. 9, pp. 1465-1470, 1968.
- WOLC, A. WHITE, I. M. S.; AVENDANO, S.; HILL, W. G. Genetic variability in residual variation of body weight and conformation scores in broiler chickens. **Poultry Science**, n.88, pp. 1156-1161, 2009.
- YANG, Y.; CHRISTENSEN, O.F.; SORENSEN, D. Analysis of genetically structured variance heterogeneity and the Box-Cox transformation. In: MEETING OF THE EUROPEAN ASSOCIATION OF ANIMAL PRODUCTION, 60, 2009, Barcelona. Book of Abstracts... Barcelona: WAP, 2009, p.598.
- ZAMPAR, A. **Comparação de estimativas de diferenças esperadas de progênie por três metodologias em touros da raça Nelore**. 2007. Dissertação (Mestrado em Qualidade e Produtividade Animal) - Faculdade de Zootecnia e Engenharia de Alimentos (FZEA), Universidade de São Paulo, Pirassununga, 2007.
- ZETTELER, P.E.; KNOL, E.F. Possibilities for selection for uniformity in pig carcasses. In: 56th EAAP Meeting, Uppsala. **Proceedings**...Uppsala: EAAP, 2005.

## **CAPÍTULO 2. SELEÇÃO CONTEMPLANDO UNIFORMIDADE NO GANHO DE PESO DO NASCIMENTO À DESMAMA EM BOVINOS NELORE: ESTUDOS PRELIMINARES**

**RESUMO** - Estudos preliminares foram desenvolvidos com o objetivo de comparar três modelos, implementados em dois passos, para estimação de variância aditiva sobre a dispersão residual do ganho de peso do nascimento à desmama (GND) de bovinos Nelore. No primeiro passo foram estimados, sob um modelo animal, componentes de variância para a média de GND, sendo obtidas as soluções para os diferentes efeitos considerados, permitindo a predição dos resíduos ( $\hat{e}$ ) associados a cada observação. Seis medidas de dispersão residual ( $d_i$ ) foram obtidas para a progênie de cada touro e utilizadas para estimação de componentes de variância sob a metodologia de máxima verossimilhança restrita. Os modelos 1 e 2 contemplaram, respectivamente, efeito aditivo do touro e efeito aditivo de pai do touro sobre  $d_i$ . O modelo 3 foi caracterizado pela análise, sob modelo touro, de uma variável binária ( $Y$ ), definida com base na magnitude dos resíduos preditos de cada observação considerando-se um intervalo ótimo para este critério. O modelo 2 foi considerado superior ao modelo 1 para estudar a variabilidade genética para dispersão residual com base em medidas de dispersão na progênie de touros. Encontrou-se razoável evidência de variabilidade genética os resíduos de GND expressos em escala binária (modelo 3).

**Palavras-chave:** dispersão residual, modelo touro, modelo de limiar, variabilidade

## 2.1. INTRODUÇÃO

Estudos iniciais foram desenvolvidos com o objetivo de investigar a viabilidade de se considerar diferentes estimadores de dispersão residual da progênie de touros Nelore como meio de selecionar para maior uniformidade no ganho de peso do nascimento à desmama (GND), o que se justificaria sobretudo pela menor complexidade e maior viabilidade prática desta abordagem quando comparada ao uso de modelos contemplando heterogeneidade genética sobre a variância residual num único passo.

O emprego de seleção para maior uniformidade em GND se justificaria em virtude de uma relação não-linear entre o peso e o preço de comercialização dos bezerros (que constitui a principal unidade de comercialização em rebanhos de cria do Brasil Central). Além disso, bezerros leves não são desejados (inclusive pela associação do peso à desmama com pesos em idades posteriores), assim como bezerros muito pesados também não são desejáveis em alguns sistemas de produção, uma vez que isto poderia implicar em resposta correlacionada sobre a produção de leite e, conseqüentemente, sobre as exigências nutricionais das vacas.

A estimação de parâmetros genéticos é crucial para a obtenção de predições acuradas quanto ao valor genético dos animais candidatos à seleção, para uma dada característica numa população, sendo que, dada a escassez de informações na literatura quanto à variabilidade genética entre touros para dispersão do GND de sua progênie, este estudo preliminar concentrou-se na identificação de modelos que permitissem a estimação destes parâmetros e na investigação quanto à confiabilidade das estimativas obtidas.

Investigou-se o uso de estimadores com base na dispersão de efeitos residuais preditos para GND, como uma extensão da abordagem proposta por FRIES et al. (2006). Entende-se que a presente abordagem também seria próxima das empregadas nos estudos relacionados à heterogeneidade genética sobre a variância residual.

Em termos biológicos, o valor genético predito com base na dispersão residual poderia estar associado à maior ou menor sensibilidade dos genótipos à variação ambiental, abrindo assim a oportunidade de seleção contemplando homogeneidade de produto final (SANCRISTÓBAL-GAUDY et al., 1998; SORENSEN & WAAGENPETERSEN, 2003).

Dentre os estimadores de dispersão que poderiam ser empregados neste estudo, investigou-se a possibilidade do uso de *Qn scale* e *Sn scale* (ROUSSEEUW &

CROUX,1993) por se tratarem de estimadores robustos de dispersão com propriedades interessantes para a situação em estudo (não pressupõem distribuição simétrica das observações), ao passo que medidas de conceito próximo à variância residual e seu logaritmo natural têm sido usadas para estimação de componentes de variância (VC) sob modelos considerando heterogeneidade genética sobre a variância residual de características de produção em outras espécies animais (e.g. SANCRISTÓBAL-GAUDY, 1998; MULDER et al. 2007), o que seria conveniente para fornecer uma base de comparação com os resultados obtidos neste estudo.

Além disso, o estudo quanto à variabilidade da progênie de touros considerando medidas robustas de dispersão (MAD, *Sn scale* e *Qn scale*) permitiria avaliar se o uso de estimadores robustos poderia minimizar o viés na predição do valor genético para variabilidade da progênie de touros com pequeno número de filhos e/ou com progênie em grupos de contemporâneos “mais variáveis”, que seria esperada ao se usar estimadores não robustos de dispersão (HILL, 2004; FRIES et al., 2006).

Isto se justificaria pelo fato de que, no caso de seleção por truncamento para uma característica com distribuição normal, numa situação em que indivíduos distribuídos entre grupos com mesma média para uma dada característica fossem candidatos à seleção, seria esperado que, à medida que a seleção se tornasse mais intensa, os indivíduos dos grupos com maior dispersão teriam maior probabilidade de serem selecionados (HILL, 2004).

Outra possível estratégia de seleção para uniformidade de produção seria a transformação dos resíduos preditos para uma escala categórica visando expressar maior ou menor uniformidade, como proposto por KANIS et al. (2006). Segundo estes autores, a utilização deste procedimento permitiria considerar a variação como uma característica individual, ao passo que no caso de medidas de dispersão da progênie, o número de observações disponíveis se limitaria ao número de touros com mais de um produto no arquivo de dados.

Este estudo foi desenvolvido como meio de investigar diferentes abordagens para seleção contemplando uniformidade no ganho de peso do nascimento à desmama em bovinos Nelore, considerando-se o emprego de estimadores de dispersão residual da progênie de touros e a transformação de resíduos preditos para uma escala categórica.

## 2.2.MATERIAL E MÉTODOS

No primeiro passo, utilizaram-se os dados de ganho de peso do nascimento à desmama (GND) de 103.748 animais da raça Nelore, provenientes do banco de dados do programa PAINT. Nas análises, foram incluídos apenas os animais provenientes de inseminação artificial (IA) ou monta controlada, para que fosse possível a identificação correta de seus pais. Os dados de GND foram pré-ajustados para os efeitos de idade da vaca ao parto (idade padrão de 7 anos), idade do bezerro (idade padrão de 205 dias) e data juliana média dentro da estação de nascimento, sendo as observações resultantes referidas como ganho do nascimento à desmama pré-ajustado (GNDa). Foram utilizados fatores de ajuste previamente estimados com dados da base Aliança Nelore, da qual o conjunto de dados utilizado neste estudo faz parte. Os procedimentos de consistência foram realizados por meio do *software* SAS (SAS INSTITUTE, 2002).

Após a verificação da consistência dos dados, procedeu-se a estimação de componentes de variância, através do método de máxima verossimilhança restrita (REML), seguida da predição do valor genético (EBV) de cada animal para GNDa, aplicando-se o *software* MTDREML (BOLDMAN et al., 1995), usando o modelo animal completo:

$$\mathbf{GNDa} = \mathbf{Xgc} + \mathbf{Z}_1\mathbf{ad} + \mathbf{Z}_2\mathbf{am} + \mathbf{Z}_3\mathbf{ap} + \mathbf{e},$$

considerando o efeito fixo de grupo de contemporâneos (**gc**, determinado pela concatenação das informações de rebanho, ano, estação, sexo, data de pesagem e grupo de manejo) e os efeitos aleatórios aditivo direto (**ad**), aditivo materno (**am**), de ambiente permanente devido à mãe (**ap**) e residual (**e**), com as respectivas matrizes de incidência de cada efeito sobre GND ( $X$ ,  $Z_1$ ,  $Z_2$  e  $Z_3$ ).

Caracterizando as pressuposições deste modelo, as matrizes de esperança e (co)variância dos efeitos considerados são dadas, respectivamente, por:

$$E(\mathbf{GND}) = \mathbf{Xgc}; \quad E(\mathbf{ad}) = E(\mathbf{am}) = E(\mathbf{ap}) = E(\mathbf{e}) = 0$$

$$e \quad \text{VAR} \begin{pmatrix} \mathbf{ad} \\ \mathbf{am} \\ \mathbf{ap} \\ \mathbf{e} \end{pmatrix} = \begin{pmatrix} A\sigma_{ad}^2 & A\sigma_{dm} & 0 & 0 \\ A\sigma_{dm} & A\sigma_{am}^2 & 0 & 0 \\ 0 & 0 & I\sigma_{ap}^2 & 0 \\ 0 & 0 & 0 & I\sigma_e^2 \end{pmatrix},$$

onde  $\sigma_{ad}^2$  é a variância genética aditiva para efeito direto,  $\sigma_{am}^2$  é a variância genética aditiva para efeito materno,  $\sigma_{dm}$  é a covariância entre os efeitos aditivos direto e materno,  $\sigma_{ap}^2$  é a variância do efeito de ambiente permanente devido à mãe e  $\sigma_e^2$  é a variância residual,  $\mathbf{A}$  é a matriz de parentesco entre os indivíduos e  $\mathbf{I}$  é uma matriz identidade de ordem igual ao número de observações.

Para estimação de componentes de variância (VC) por REML, adotou-se variância menor do que  $10^{-9}$  para a função simplex como critério de convergência.

Tendo em vista que mais de 80% das vacas com informação de desempenho de progênie não possuíam informação de desempenho próprio para a característica em estudo, optou-se por também assumir covariância nula entre efeito aditivo direto e aditivo materno ( $\sigma_{dm} = 0$ ), uma vez que, sob esta estrutura de *pedigree*, seriam esperadas estimativas bastante viesadas desta covariância (HEYDARPOUR et al.; 2008). Os demais componentes de covariância entre os efeitos aleatórios foram assumidos iguais a zero.

No segundo passo, para cada produto, utilizaram-se as soluções obtidas para os efeitos de gc, ad, am e ap para se calcular o resíduo RES1, obtido com base no ajuste do valor fenotípico de GNDa para efeito fixo de GC, efeitos aditivos associados à vaca (aditivo direto e materno) e efeitos aleatórios de ambiente permanente, permitindo a obtenção de um resíduo reunindo efeitos aditivo de touro, segregação mendeliana e efeito aleatório residual. A princípio, tal definição de resíduo permitiria contemplar, simultaneamente, a ação de efeitos aditivos de touro sobre a variância residual bem como efeitos de prepotência dos touros, como sugerido por HOHENBOKEN (1985).

Para cada touro, foram calculadas seis diferentes medidas de dispersão para as observações RES1 de sua progênie: i) Desvio-padrão (SD); ii) *Median absolute deviation* (MAD); iii) *Sn scale* (Sn); iv) *Qn scale* (Qn); v) Variância (VAR) e vi) Logaritmo natural de VAR (ln).

*Estimadores de dispersão residual para incremento de uniformidade considerando efeito aditivo de touro (Modelo 1)*

Estimaram-se componentes de variância (VC) para cada medida de dispersão das observações de RES1 na progênie, sob um modelo associando cada medida de dispersão

ao efeito aditivo de touro (Modelo 1). Este modelo foi implementado por meio do *software* MTDREML e pode ser descrito pela equação:

$$d = Zs + \varepsilon, \text{ (Modelo 1),}$$

em que **d** é o vetor de medidas de dispersão residual da progênie dos touros, **s** é o vetor de efeito aditivo de touro, **ε** é o vetor de efeitos residuais (no segundo passo) e **Z** é a matriz de incidência que relaciona os elementos de **s** a **d**.

As matrizes de esperança e (co)variância do **Modelo 1** são dadas por:

$E(\mathbf{d}) = \mu$  (i.e. dispersão residual média),  $E(\mathbf{s}) = E(\boldsymbol{\varepsilon}) = 0$ ; Variância  $(\mathbf{d}) = A\sigma_s^2 + I\sigma_\varepsilon^2$ , onde **A** é a matriz de parentesco entre os touros (baseada nos pais e mães de touros) e **I** é uma matriz identidade. Neste modelo assumiu-se **s** e **ε** normalmente distribuídos, i.e.  $\mathbf{s} \sim N(0, A\sigma_s^2)$  e  $\boldsymbol{\varepsilon} \sim N(0, I\sigma_\varepsilon^2)$ , e que  $\sigma_s^2$  equivale a 1/4 da variância aditiva para dispersão residual, uma vez que as observações de resíduo obtidas no primeiro passo das análises referem-se à progênie dos touros.

Para as análises de medidas de dispersão, foram considerados apenas touros com, no mínimo, cinco produtos. Em virtude desta restrição, as análises foram conduzidas com dados de dispersão de 740 famílias de touros, admitindo-se que uma variância da função simplex menor que  $10^{-12}$  indicaria convergência.

Com base na estimativa de VC pelo Modelo 1, calculou-se a herdabilidade ( $h^2$ ) de cada medida de dispersão, segundo a fórmula  $h^2 = \sigma_a^2 / \sigma_d^2$ , em que  $\sigma_a^2$  é a variância aditiva para dispersão residual e  $\sigma_d^2$  é a variância “fenotípica” para dispersão residual na progênie. Uma medida de conceito similar à *evolubility* (HOULE, 1992), o coeficiente de variação genético aditivo (Ev), foi calculada como a razão entre o desvio-padrão genético aditivo e dispersão média, para cada medida. As estimativas de  $h^2$  e Ev serviram de base para comparação entre as diferentes medidas de dispersão e modelos considerados.

*Estimadores de dispersão residual para incremento de uniformidade considerando efeito aditivo de pai de touro (Modelo 2)*

Com base em indícios de que o problema de decomposição de medidas de dispersão da progênie entre efeito de touro e residual não fosse tratável sob o modelo 1 (dado o uso de uma única medida de dispersão por touro e a matriz de parentesco disponível), investigou-se o uso de um modelo alternativo para analisar as medidas de

dispersão residual, associando o vetor de dados de dispersão ao efeito aleatório de pai dos touros e usando a informação de parentesco disponível (i.e. basicamente avós paternos dos touros).

Utilizou-se o mesmo procedimento em dois passos para estimação de VC de medidas de dispersão do resíduo RES1 da progênie de touros descrito anteriormente, porém decompondo-se cada medida de dispersão da progênie dos touros entre efeito aditivo de pai do touro e um efeito residual aleatório (Modelo 2), por meio do software MTDFREML. O Modelo 2 pode ser descrito por:

$$d = Z_a g + \varepsilon, \text{ (Modelo 2),}$$

em que  $\mathbf{d}$  é o vetor de medidas de dispersão residual da progênie dos touros,  $\boldsymbol{\mu}$  é a dispersão residual média,  $\mathbf{g}$  é o vetor de efeito aditivo de pai de touro e  $\boldsymbol{\varepsilon}$  é o vetor de efeitos residuais (no segundo passo) e  $\mathbf{Z}_a$  é a matriz de incidência que relaciona os elementos de  $\mathbf{g}$  a  $\mathbf{d}$ .

As matrizes de esperança e (co)variância do **Modelo 2** são dadas por:

$E(\mathbf{d}) = \boldsymbol{\mu}$  (i.e. dispersão residual média),  $E(\mathbf{g}) = E(\boldsymbol{\varepsilon}) = 0$ ; Variância  $(\mathbf{d}) = \mathbf{A}\sigma_g^2 + \mathbf{I}\sigma_\varepsilon^2$ , onde  $\mathbf{A}$  é a matriz de parentesco entre os pais de touros (baseada nos avôs paternos dos touros) e  $\mathbf{I}$  é uma matriz identidade. Neste modelo assume-se  $\mathbf{g}$  e  $\boldsymbol{\varepsilon}$  normalmente distribuídos, i.e.  $\mathbf{g} \sim N(0, \mathbf{A}\sigma_g^2)$  e  $\boldsymbol{\varepsilon} \sim N(0, \mathbf{I}\sigma_\varepsilon^2)$ , e que  $\sigma_g^2$  equivale **1/16** da variância aditiva para dispersão residual, uma vez que as predições de resíduo no primeiro passo das análises referem-se aos netos dos animais aos quais o efeito  $\mathbf{g}$  é associado.

Como meio de verificar a confiabilidade e/ou confirmar as estimativas obtidas pela metodologia REML, investigou-se o uso de duas abordagens: validação cruzada e inferência bayesiana. O procedimento de validação cruzada foi implementado por meio de uma adaptação da metodologia de *data splitting* usada por PEREZ-ENCISO et al. (1993), a fim de verificar a habilidade preditiva dos modelos considerando diferentes medidas de dispersão e estimação de VC por REML. Seguindo esta metodologia, para cada modelo, os conjuntos de dados foram divididos aleatoriamente em duas amostras (A e B) com igual número de observações, de modo que os componentes de variância estimados em A, bem como parâmetros de posição, fossem usados para prever as observações da amostra B (e vice-versa), em 50 repetições.

Segundo esta metodologia, a habilidade preditiva dos modelos poderia ser comparada por meio da correlação entre as observações e as respectivas predições obtidas com base nos parâmetros estimados com a amostra complementar, bem como por meio do viés e erro quadrático médio (MSE).

A outra abordagem alternativa estudada foi a estimação de componentes de variância sob o Modelo 2 por meio de inferência bayesiana, com uso do software MTGSAM (VAN TASSEL & VAN VLECK, 1995). Porém, após tentativas de estimar componentes de variância para estimadores de dispersão usando-se este procedimento, foram identificados problemas de convergência, mesmo utilizando-se cadeias de tamanho considerável e diferentes *priors*, sendo este problema também observado sob o Modelo 1.

Uma possível explicação para problemas de convergência sob esta metodologia poderia residir no uso do algoritmo de *Gibbs sampling* (GS) para estimação de VC em conjuntos de dados relativamente pequenos (como é o caso dos modelos em estudo), uma vez que pequenos conjuntos de dados poderiam sobrepujar a qualidade dos geradores de números pseudo-aleatórios usados no algoritmo de GS (MISZTAL, 2008).

Com base nas estimativas REML de variância aditiva para dispersão residual, as medidas  $h^2$  e  $E_v$  foram calculadas conforme as definições apresentadas anteriormente para o Modelo 1.

### *Análise de resíduos em escala categórica (Modelo 3)*

Foram utilizados os mesmos dados de GND descritos no item anterior (informação de ganho de peso do nascimento à desmama de 103.748 animais), também por meio de um procedimento em dois passos.

Resíduos (RES2) foram preditos com base no ajuste das observações de GNDa para as soluções referentes aos efeitos aditivos de touro e vaca (direto e materno), aleatório de ambiente permanente e fixo de GC obtidas no primeiro passo, conforme proposto no estudo de KANIS et al. (2006), sendo calculados a média (MR) e desvio-padrão (SR) dos valores de RES2.

Para cada produto, observações de RES2 foram transformados numa variável binária (Y), com base num intervalo “ótimo”, i.e valor 1, caso RES2 estivesse dentro do intervalo compreendido entre  $MR \pm 1*SR$ , ou valor igual a 0, caso contrário. Conforme

esperado, aproximadamente 68% das observações apresentaram valor 1 para Y (Teorema do Limite Central).

MISZTAL et al. (1989) relataram a possibilidade de problemas nas análises de variáveis categóricas no caso da existência de níveis de efeitos fixos com uma única categoria (*extreme category problem*), sendo que, visando contornar este eventual problema, foram desconsiderados nas análises os grupos de contemporâneos (GCs) com o mesmo valor para todas as observações de Y, de modo que no arquivo final foram mantidas 98.583 observações, distribuídas em 4804 GCs e referentes aos filhos de 739 touros.

As distribuições *a posteriori* dos componentes de variância e da herdabilidade de Y foram obtidas com o uso do software MTGSAM Threshold (VAN TASSEL et al.,1998), considerando um modelo touro incluindo efeitos aditivo de touro, de GC e efeito aleatório residual ( $\epsilon$ ). Como a variância residual ( $\sigma^2_\epsilon$ ) não é estimável sob modelos de limiar, um valor arbitrário teria que ser atribuído para a mesma, tendo sido utilizada a parametrização  $\sigma^2_\epsilon = 1$ , conforme sugerido por GIANOLA & FOULLEY (1983). Admitiu-se a variância de touro como sendo  $\frac{1}{4}$  da variância aditiva.

O comprimento das cadeias de Markov utilizadas para estimativa de herdabilidade de Y foi de 50.000 ciclos, com *burn-in* dos 4.000 ciclos iniciais, tendo sido atribuído valor não informativo como *prior para* a variância do efeito de touro. A análise de convergência e de correlação serial foi realizada por meio do pacote *Bayesian Output Analysis* (SMITH, 2005) do *software* R.

As médias *a posteriori* das soluções para efeito de touro na escala subjacente foram utilizadas para cálculo das DEPs dos touros para porcentagem de produtos dentro do intervalo desejado ( $Y=1$ ), por meio da transformação das soluções para a escala de porcentagem, expressando-as como um desvio de 50%, conforme procedimento empregado no estudo de ELER et al. (2002).

### 2.3. RESULTADOS E DISCUSSÃO

#### *Primeiro passo das análises*

Analisando-se as estatísticas descritivas referentes às observações obtidas após a verificação de consistência dos dados utilizados no primeiro passo das análises (Tabela 2), pode-se perceber que estes valores foram próximos daqueles que têm sido relatados para esta característica na raça Nelore (e.g. ORTIZ PEÑA et al., 2004; MARCONDES et al. 2000). Os valores de assimetria e curtose observados permitem inferir que a distribuição de GNDa se aproxima da normalidade, uma vez que a estimação de componentes de variância com o *software* MTDFREML pressupõe distribuição normal dos efeitos aleatórios.

Tabela 2. Estatísticas referentes aos dados de ganho de peso do nascimento a desmama de bovinos Nelore

Variável	Nº de observações	Valores expressos em kg						Assimetria	Curtose
		Média	Mediana	DP	CV	Mínimo	Máximo		
GNDa	103.748	169,40	167,95	39,67	23,4%	51,11	297,83	0,188	-0,212

DP=desvio-padrão, em kg. CV=coeficiente de variação.

As estimativas de herdabilidade obtidas neste primeiro passo foram de  $0,16 \pm 0,014$  para efeito direto e  $0,12 \pm 0,020$  para efeito materno sobre GNDa, estando dentro do que seria esperado para esta característica em populações Nelore. A variância residual estimada (VE) foi de 198,92 kg<sup>2</sup>.

Na Tabela 3, apresentam-se as estatísticas descritivas referentes aos dados de dispersão de resíduo (RES1) da progênie considerados neste estudo. Com base nos valores de desvio-padrão para os diferentes estimadores, pode-se verificar as diferenças de escala entre os mesmos.

#### *Segundo passo das análises*

Executaram-se análises incluindo informação de parentesco entre os touros (Modelo 1), por meio do desenvolvimento de um programa para recuperação de informação dos pais de touros com base num arquivo específico de *pedigree*.

Tabela 3. Estatísticas descritivas\* referentes aos dados de medidas de dispersão do resíduo das progênie de touros: desvio-padrão (SD), variância (VAR), *median absolute deviation* (MAD), *Sn scale* (Sn), *Qn scale* (Qn) e log de VAR (ln).

Estimativa	N	MIN	MAX	Média	DP	CVd(%)
PROG	740	5	5964	140	468	-
SD	740	5,50	58,93	20,06	4,94	24,61
VAR	740	30,3	3472,9	426,7	235,0	55,1
MAD	740	1,97	39,12	12,07	3,65	30,26
Sn	740	2,35	57,26	18,42	5,17	28,08
Qn	740	3,80	90,06	21,57	6,76	31,36
ln	740	3,41	8,15	5,94	0,50	8,43

\*PROG=tamanho de progênie, N=número de observações, MIN=mínimo, MAX=máximo, DP=desvio-padrão e CVd=coeficiente de variação para dispersão dentro de touro, em % .

Como é possível observar na tabela 4, dada a magnitude dos erros-padrão da herdabilidade, todas estimativas de  $h^2$  para medidas de dispersão da progênie não foram estatisticamente diferentes de zero sob o Modelo 1, tomando-se como base um intervalo de confiança de 95%. De modo geral, estes resultados geraram dúvidas quanto à confiabilidade deste modelo para estudos sobre variabilidade de GND, inclusive por sua sensibilidade ao emprego de diferentes *priors* para variância de touro.

A estrutura dos dados para análise de componentes de variância usando medidas de dispersão da progênie implica em apenas uma observação (medida de dispersão) para cada touro, sob o Modelo 1. A menos que os componentes de variância para dispersão residual fossem conhecidos previamente, os efeitos de touro e residual só poderiam ser discriminados adequadamente por meio da informação de parentesco entre touros e, neste caso, todas as medidas de dispersão (exceto Qn) apresentaram variância aditiva muito próxima de zero.

Frente a estas questões, procurou-se investigar alternativas que permitissem verificar a confiabilidade das estimativas obtidas. Inicialmente, procurou-se estimar VC segundo o modelo 1 por meio de inferência bayesiana, considerando a informação de parentesco disponível, com uso do *software* MTGSAM (VAN TASSEL & VAN VLECK, 1995). Porém, as estimativas de variância aditiva sob esta metodologia mostraram-se sempre crescentes, não sendo possível identificar-se convergência, reforçando o problema de estimabilidade dos parâmetros de interesse. Nestas análises, foram geradas cadeias com comprimento de  $10^6$  ciclos, com período de descarte dos primeiros 10.000 ciclos e uso de *priors* não informativos.

Tabela 4. Estimativas de variâncias aditiva ( $va$ ) e residual ( $ve$ ), herdabilidade ( $h^2$ ), erro padrão de  $h^2$  ( $ep$ ) e *evolvability* ( $Ev$ ) para medidas de dispersão de resíduo da progênie, utilizando-se o Modelo 1

Medida de dispersão	$va$	$ve$	$h^2$	$ep$	$Ev(\%)$
SD	0,000	24,34	0,00	0,54	0,04
VAR	0,040	50000	0,00	0,576	0,05
MAD	0,000	13,31	0,00	0,56	0,00
<i>Sn scale</i>	0,000	26,71	0,00	0,54	0,00
<i>Qn scale</i>	8,907	36,91	0,19	0,700	13,84
<i>ln</i>	0,000	0,2502	0,00	0,504	0,00

O uso do Modelo 2 (consideração de efeito aditivo de pai de touro) para estimar os componentes de variância para as medidas de dispersão do resíduo da progênie mostrou-se mais estável do que o Modelo 1, de modo que as estimativas por REML com a inclusão ou não da informação de parentesco mantiveram-se próximas (dados não apresentados). Os valores estimados para  $h^2v$  entretanto, apresentaram magnitude discrepante entre diferentes estimadores. Além disso, os erros-padrão foram elevados, indicando que a consideração deste parâmetro poderia não ser adequada para se prever, de modo acurado, a resposta à seleção sob esta abordagem (Tabela 5).

Um fato que chama a atenção é que, dentre as medidas tidas como robustas, apenas para *Qn scale* foi identificada variância aditiva não-nula e sua estimativa de  $h^2$  foi a maior dentre as medidas de dispersão do resíduo consideradas (Tabela 5).

De modo geral, os valores estimados para *evolvability* (Tabela 5) foram inferiores àqueles relatados para características de produção (entre 25 e 60%, conforme revisto no estudo de MULDER et al., 2007). Diferenças de escala entre estimadores também podem ter influenciado este resultado, todavia menor *evolvability* também foi estimada para VAR, que se refere a um estimador na mesma escala daqueles considerados nos estudos anteriores.

As estimativas de herdabilidade para medidas de dispersão residual da progênie de touros não parecem estar de acordo com as de outros estudos com abordagem similar. Estudos considerando medidas para dispersão do peso ao nascer dentro de leitegada em suínos (vista como característica materna) têm encontrado baixas herdabilidades ao analisar dispersão da progênie como uma característica *per se*, porém com a ressalva de

que estes estudos consideraram a dispersão dos pesos medidos ao invés da dispersão de resíduos preditos.

WITTENBURG et al. (2008) relataram herdabilidades entre 7% a 11% para a variabilidade do peso ao nascer dentro de leitegada, usando medidas de dispersão de conceito próximo àquelas definidas como SD, VAR e LN no presente estudo. DAMGAARD et al. (2003) também encontraram herdabilidade de magnitude semelhante (0,08) para o desvio-padrão do peso ao nascer dentro de leitegada em suínos.

Tabela 5. Estimativas de variância aditiva (va) e residual (ve), herdabilidade ( $h^2$ ), erro-padrão de  $h^2$  (EP) e *evolvability* (Ev) obtidas por metodologia REML para diferentes medidas de dispersão da progênie, utilizando-se o Modelo 2

Medida dispersão	va	ve	$h^2$	EP	Ev(%)
MAD	0	9,641	0	0,144	0,00
SD	5,342	16,241	0,322	0,144	11,44
VAR	10198	31869	0,314	0,14	23,73
ln	0,056	0,158	0,347	0,148	3,96
Sn	0	18,820	0	0,120	0,00
Qn	14,168	24,972	0,548	0,148	17,99

Uma possível explicação para os resultados obtidos no presente estudo poderia ser a diferente definição de resíduo (RES1) empregada, o que poderia ter viesado as estimativas de variância aditiva relacionadas a dispersão residual. Além disso, a não consideração de efeitos ambientais no Modelo 2 também pode ter contribuído para que a variância aditiva tenha sido superestimada. Além disso, a grande disparidade entre estimativas de  $h^2$  para dispersão residual sob diferentes modelos (Modelo 1 vs. Modelo 2) indicaria que o uso deste parâmetro como medida de variabilidade genética talvez não fosse o mais adequado para esta situação, podendo-se empregar a *evolvability* para este fim, conforme sugerido por MULDER et al. (2007).

Como resultado do procedimento de validação cruzada empregado, percebe-se que a correlação média entre a dispersão observada e aquela predita com base nos parâmetros estimados (usando-se a amostra complementar) foi ligeiramente inferior para as medidas robustas Sn e MAD, ao passo que o menor viés foi observado no caso de ln e MAD. Além disso, os erros-padrão do viés foram de magnitude elevada, inclusive para estas últimas medidas (Tabela 6).

De modo geral, não foram perceptíveis grandes diferenças entre a habilidade preditiva dos modelos considerando diferentes medidas de dispersão. As diferenças de escala entre as medidas de dispersão, claramente evidenciada nos resultados referentes ao erro quadrático médio (MSE), tornam difícil a interpretação da habilidade preditiva de cada medida. No caso de desempenho preditivo semelhante, com o mesmo número de parâmetros, poderia se assumir que os modelos mais promissores para seleção contemplando uniformidade seriam aqueles para os quais foi estimada maior *evolubility* (i.e. VAR e Qn).

Tabela 6. Estatísticas\* referentes ao procedimento de validação cruzada, para as diferentes medidas de dispersão do resíduo da progênie

Médias (Erro-Padrão) por estimador de dispersão						
Estatística	In		MAD		Qn	
MSE	0,2113	(0,0031)	11,8852	(0,1472)	31,9642	(0,5013)
viés	-0,0006	(0,0057)	-0,0006	(0,0497)	-0,0197	(0,0673)
r	0,5747	(0,0069)	0,5295	(0,0055)	0,5746	(0,0063)
Estatística	SD		Sn		VAR	
MSE	19,3342	(0,2334)	22,7329	(0,2595)	36185	(566)
viés	-0,0059	(0,0548)	-0,0026	(0,0668)	-0,3118	(2,3199)
r	0,5711	(0,0061)	0,5243	(0,0061)	0,5763	(0,006)

\*MSE=erro quadrático médio, r=correlação média entre dispersão observada e predita.

Após a análise *bayesiana* da variável categórica Y, os resultados dos testes de Geweke e de Heidelberger & Welch deram suporte para indício de convergência. A obtenção de um baixo valor de correlação serial para a cadeia de Gibbs (-0,0172) não evidenciou problema de redundância de informações.

Como a distribuição *a posteriori* das estimativas de herdabilidade apresentou-se, aproximadamente, simétrica, a média da herdabilidade obtida ao analisarem-se os resíduos preditos para GNDa como uma variável categórica (Y) poderia ser usada como uma medida de tendência central para  $h^2$ , sendo que seu valor, calculado com base em 914 sub-amostras, foi igual 0,1579 (com valores iguais a 0,1341 e 0,1839 para os limites inferior e superior da região de credibilidade com 95% de probabilidade, respectivamente) (Figura 1).

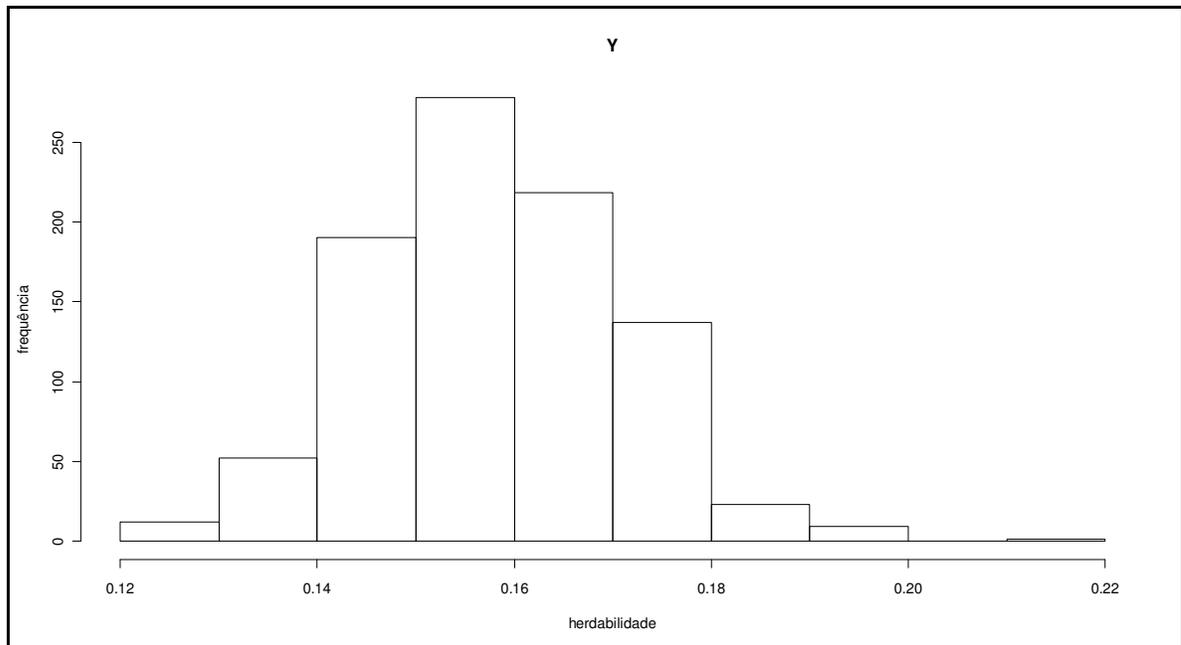


Figura 1. Histograma da distribuição *a posteriori* da herdabilidade ( $h^2$ ) para os resíduos de GNDa expressos em escala categórica (Y).

Apesar da herdabilidade baixa a moderada para Y, os valores encontrados foram superiores aos obtidos por KANIS et al. (2006) ao utilizarem abordagem similar para análise da variabilidade no peso de abate e no peso de carcaça quente de suínos. Os valores estimados no presente estudo sugeririam existência de variabilidade genética entre touros para a variância residual e a possibilidade de obtenção de resposta moderada à seleção para aumento da proporção de produtos dentro de um intervalo considerado ótimo.

KANIS et al. (2006) relataram que a consideração de efeitos fixos na análise de uma característica semelhante a Y, levou à fixação das estimativas de  $h^2$  no limite deste parâmetro, não fornecendo resultados interpretáveis. Tal ocorrência pode se dever a utilização de um modelo linear para análise desta característica binária, apesar dos autores não terem explicitado claramente esta informação.

No presente estudo, mesmo após a consideração do efeito de GC sobre Y, obteve-se convergência com razoável evidência de herdabilidade não-nula para esta característica. Este tipo de modelo, comparado ao Modelo 2, teria a vantagem de permitir a remoção, em algum grau, de efeitos ambientais presentes nos resíduos de cada observação que podem ter viesado as estimativas de variância aditiva obtidas sob o Modelo 2.

Uma regressão linear da média das soluções *a posteriori* para efeito de touro ( $DEP_Y$ , transformadas em %) sobre a porcentagem de produtos dentro do intervalo desejado (proporção de produtos com  $Y=1$ ) explicaria, aproximadamente, 78% da variação observada para porcentagem de produtos com resíduo dentro do intervalo “ótimo” (Figura 2a).

Percebe-se uma relação não linear entre DEP para média de GNDa ( $DEP_m$ ) e  $DEP_Y$  (Figura 2b). Nesta situação, a seleção dos touros com as  $DEP_m$  mais elevadas poderia proporcionar resposta desfavorável em termos da proporção de produtos dentro do intervalo ótimo (decorrente da maior probabilidade de gerar produtos com GNDa elevado). Todavia, haveria oportunidade de seleção para incremento da média e uniformidade, por meio do uso de índices combinando  $DEP_m$  e  $DEP_Y$ .

Dada a amplitude das DEPs preditas para  $Y$ , seria esperada uma diferença de mais de 25 unidades percentuais na probabilidade de produtos com GNDa dentro do intervalo “ótimo”, comparando-se os melhores e os piores touros para este critério e assumindo-se acasalamento com vacas de mérito médio para  $Y$ .

Apesar da definição de  $Y$  ter sido feita com base num critério arbitrário (que influenciou diretamente a média de  $Y$  e a proporção de observações por categoria e, portanto, os parâmetros estimados), os resultados obtidos indicaram a possibilidade de se utilizar abordagens similares, no caso de situações com bonificação estabelecida para produtos com peso dentro de um intervalo ótimo, dada a evidência de variabilidade genética entre touros para este tipo de característica.

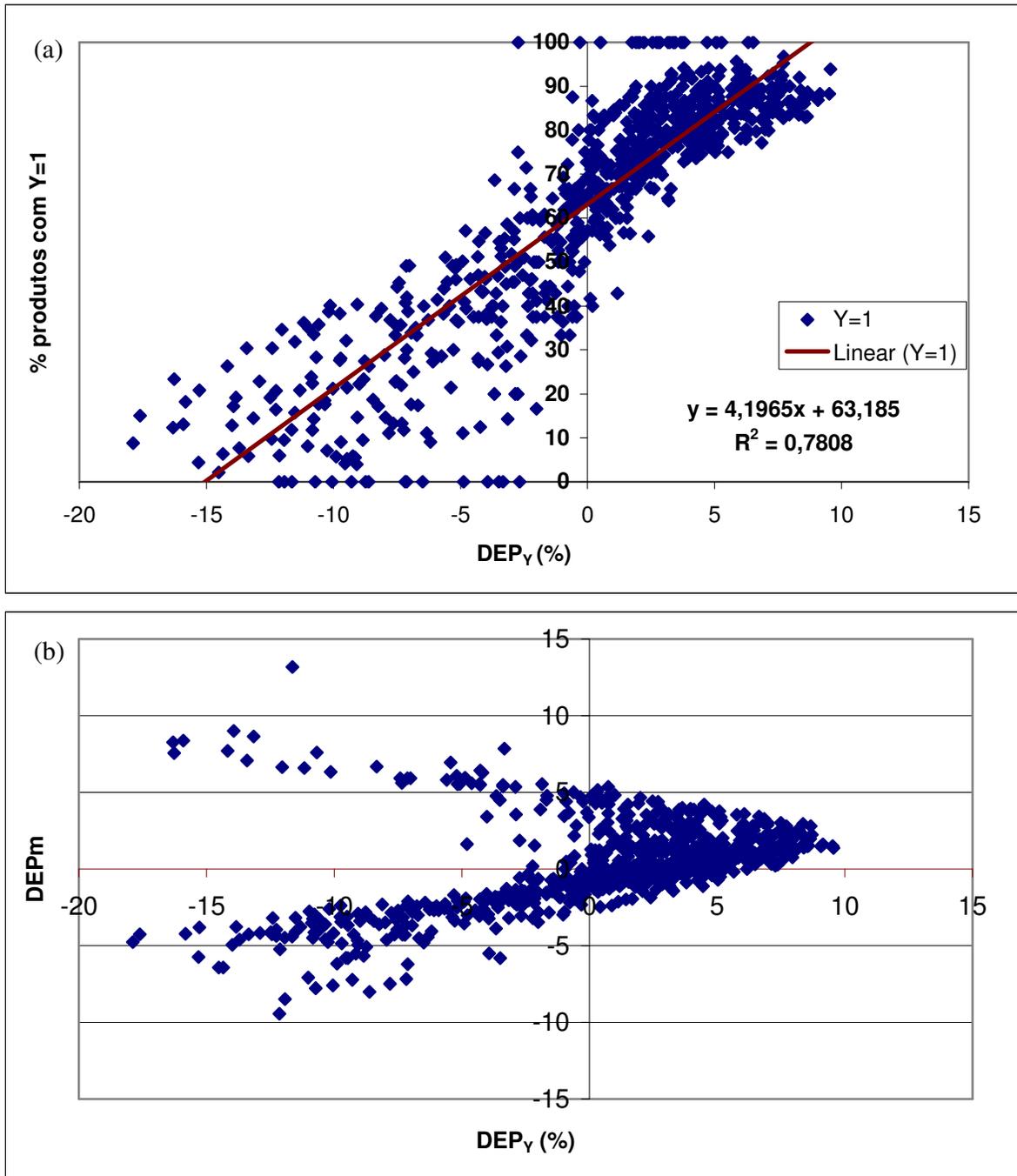


Figura 2: a) Relação entre a média *a posteriori* dos efeitos de touro ( $DEP_Y$ , em %) para a variável  $Y$  e a porcentagem (%) de produtos com resíduo dentro do intervalo "ótimo" ( $Y=1$ ). b) Relação entre  $DEP_Y$  e a DEP para a média de GNDa ( $DEP_m$ )

## 2.4. CONCLUSÕES

O uso do Modelo 1 não se mostrou adequado para o estudo sobre variabilidade genética para dispersão residual, no caso dos dados considerados no presente estudo.

A associação de medidas de dispersão na progênie ao efeito de pai de touro (Modelo 2) possibilitou a obtenção de estimativas da variância aditiva sobre a dispersão residual com razoável evidência de convergência. Futuros estudos devem ser conduzidos para investigar a aplicação deste modelo usando dados de dispersão dos resíduos preditos (livres do efeito de touro e segregação mendeliana), de modo que os resultados deste estudo possam ser mais comparáveis com aqueles que consideraram heterogeneidade genética sobre a variância residual.

Encontrou-se razoável evidência de variabilidade genética entre touros para os resíduos expressos em escala binária (Modelo 3), indicando que a seleção com base neste critério poderia permitir algum incremento na uniformidade de GND. Seria indicado o estudo desta abordagem considerando outras características de produção em bovinos de corte, sobretudo em situações com intervalo ótimo definido.

## 2.5. REFERÊNCIAS

- BOLDMAN, K.G.; KRIESE, L.A.; VAN VLECK, L. D.; VAN TASSEL, C. P.; KACHMAN, S. D. 1995. **A manual for use for MTDFREML**: a set of programs to obtain of variance and covariances. Lincoln: Department of Agriculture /ARS. 120p.
- DAMGAARD, L.H.; RYDHMER, L.; LØVENTAL, P., GRANDINSON, K. Genetic parameters for within-litter variation in piglet birth weight and change in within-litter variation during suckling. **Journal of Animal Science**, Savoy, v. 81, pp. 604-610, 2003.
- ELER, J.P.; SILVA, J.A. II V.; FERRAZ, J.B.S., OLIVEIRA, H.N.; EVANS, J.L.; GOLDEN, B.L. Genetic evaluation of the probability of pregnancy at 14 months for Nellore heifers. **Journal of Animal Science**, Savoy, v.80, n.4, p.951-954, 2002.
- FRIES, L.A.; BRITO, F.V.; CARDOSO, V.; CARVALHEIRO, R.; ORTIZ-PEÑA, C.D.; PICCOLI, M. L.; QUEIROZ, S.A.; ROSO, V.M.; SCHENKEL, F.S.; SEVERO, J.L.P.

- Preliminary assessment of robust joint location & dispersion estimators of breeding values as applied to beef cattle traits in Brazil. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 8<sup>TH</sup>, 2006, Belo Horizonte. **Proceedings...Belo Horizonte:WCGALP**, 2006.
- GIANOLA, D.; FOULLEY, J.L. Sire evaluation for ordered categorical data with a threshold model. **Genetics Selection Evolution**, v.15, p.201-224, 1983.
- HEYDARPOUR, M.; SCHAEFFER, L.R.; YAZDI, H. Influence of population structure on estimates of direct and maternal parameters. **Journal of Animal Breeding and Genetics**, Berlin, n 125, pp. 89-99, 2008.
- HILL, W. G. Heterogeneity of genetic and environmental variance of quantitative traits. **Journal of Indian Society of Agricultural Statistics**, v. 57, pp. 49-63, 2004.
- HOHENBOKEN, W. D. The manipulation of variation in quantitative traits: a review of possible genetic strategies. **Journal of Animal Science**, Savoy, n. 60, pp. 101-110, 1985.
- HOULE, D. Comparing *evolvability* and variability of quantitative traits. **Genetics**, Bethesda, v. 130, pp. 195-204, 1992.
- KANIS, E.; VAN PELT, M. L; BONEKAMP, P. R. T.; KNOL, E. F. Is within-family variation in carcass weight of pigs heritable? In: 8<sup>TH</sup> WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 2006, Belo Horizonte. **Proceedings...Belo Horizonte:WCGALP**, 2006.
- MARCONDES, C.R.; BERGMANN, J.A.G.; ELER, J.P.; FERRAZ, J.B.S.; PEREIRA, J.C.C.; PENNA, V.M. Análise de alguns critérios de seleção para características de crescimento na raça Nelore. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 52, pp. 83-89, 2000.
- MISZTAL, I. Reliable computing in estimation of variance components. **Journal of Animal Breeding and Genetics**, n. 125, pp. 363-370, 2008.
- MISZTAL, I.; GIANOLA, D.; FOULLEY, J. L. Computing aspects of a nonlinear method of sire evaluation for categorical data. **Journal of Dairy Science**, n. 72, pp. 1557–1568, 1989.
- MULDER, H. A.; BIJMA, P.; HILL, W. G. Prediction of Breeding Values and Selection Responses With Genetic Heterogeneity of Environmental Variance. **Genetics**, Bethesda, v.175, pp. 1895-1910, 2007.

ORTIZ PEÑA, C.D.; CARVALHEIRO, R.; QUEIROZ, S.A.; FRIES, L.A. Comparison of selection criteria for pre-weaning growth traits of Nelore cattle. **Livestock Production Science**, v. 86, pp.163-167, 2004.

PEREZ-ENCISO, M.; TEMPELMAN, R. J.; GIANOLA, D. A comparison between linear and Poisson mixed models for litter size in Iberian pigs. **Livestock Production Science**, v. 35, pp. 303–316, 1993.

ROUSSEEUW, P. J.; CROUX C. Alternatives to the Median Absolute Deviation. **Journal of the American Statistical Association**, Alexandria, v. 88, n. 424, p. 1273-1283, 1993.

SANCRISTOBAL-GAUDY, M.; ELSEN, J-M.; BODIN, L.; CHEVALET, C. Prediction of the response to a selection for canalisation of a continuous trait in animal breeding. **Genetics Selection Evolution**, Les Ulis, n. 30, pp. 423–451, 1998.

SMITH, B. J. Bayesian output analysis program (BOA), Version 1.1.5. University of Iowa, Iowa City, IA, 2005. Disponível em: <http://www.public-health.uiowa.edu/BOA/>. Acesso em 03.mar.2009.

SORENSEN, D.; WAAGEPETERSEN, R. Normal linear models with genetically structured residual variance heterogeneity: a case study. **Genetics Research**, Cambridge, v. 82, pp. 207–222, 2003.

STATISTICAL ANALYSIS SYSTEM - SAS. SAS Institute Inc. 2002. SAS OnlineDoc® 9. Cary, NC: SAS Institute Inc.

VAN TASSELL, C. P., VAN VLECK, L.D. **A Manual for Use of MTGSAM**: A Set of Fortran Programs to Apply Gibbs Sampling to Animal Models for Variance Component Estimation. USDA, ARS, Clay Center, NE, 1995.

VAN TASSELL, C.P.; VAN VLECK, L.D.; GREGORY, K.E. Bayesian analysis of twinning and ovulation rates using a multiple-trait threshold model and Gibbs sampling. **Journal of Animal Science**, v.76, p.2048-2061, 1998.

WITTENBURG, D.; GUIARD, V.; TEUSCHER, F.; REINSCH, N. Comparison of statistical models to analyse the genetic effect on within-litter variance in pigs. **Animal**, v. 2, n. 11, pp 1559–1568, 2008.

### **CAPÍTULO 3 - ABORDAGEM EM DOIS PASSOS PARA ESTUDAR A VARIABILIDADE GENÉTICA PARA DISPERSÃO RESIDUAL NO GANHO DE PESO DO NASCIMENTO À DESMAMA NA RAÇA NELORE**

**RESUMO** – Este estudo foi desenvolvido com o objetivo de investigar uma abordagem em dois passos para estimação de parâmetros genéticos para a dispersão residual do ganho de peso do nascimento à desmama (GND) de bovinos Nelore. Num primeiro passo foram estimados, sob um modelo animal, componentes de variância para a média de GND, sendo obtidas as soluções para os diferentes efeitos considerados, permitindo a predição dos resíduos ( $\hat{e}$ ) associados a cada observação. No segundo passo, seis medidas de dispersão de  $\hat{e}$  ( $d_i$ ) foram obtidas para a progênie de cada touro e utilizadas para estimação de componentes de variância sob as metodologias de máxima verossimilhança restrita e análise de componentes de variância bayesiana não-conjugada, com base num modelo associando cada medida  $d_i$  a um efeito aditivo de pai do touro. Encontrou-se indícios de ação genética aditiva sobre a dispersão residual da característica em estudo, o que indicaria oportunidade para incremento em sua uniformidade por meio de seleção. É necessário que a metodologia empregada no segundo passo das análises seja validada.

**Palavras-chave:** *evolvability*, ganho de peso do nascimento à desmama , variância residual

### 3.1.INTRODUÇÃO

Em muitas situações, é interessante promover não apenas mudança na média de características de interesse econômico, mas também em sua variabilidade (e.g. bonificação para carcaças com características dentro de um intervalo ótimo). Nestes casos, a seleção poderia ser um meio efetivo para incremento na uniformidade de produção, caso houvesse diferenças genéticas entre os animais para variabilidade fenotípica.

Modelos assumindo heterogeneidade genética sobre a variância residual (VE) têm sido utilizados para investigar essa questão e foram revistos no estudo de MULDER et al. (2007). Ao contrário do pressuposto de igualdade de variância residual entre genótipos, em tais modelos considera-se um efeito aditivo ( $A_v$ ) associado à variância residual (em sua escala original ou logarítmica), geralmente assumindo-se uma distribuição normal multivariada com o efeito aditivo para a média da característica ( $A_m$ ). Segundo esta abordagem, têm sido encontradas evidências de variabilidade genética para a variância residual de características de produção em aves, suínos, ovinos e bovinos de leite.

A maioria dos modelos assumindo heterogeneidade genética da variância residual têm sido implementados num único passo, permitindo estimação simultânea de variâncias associadas aos valores genéticos para média ( $A_m$ ) e para variância residual ( $A_v$ ), bem como da covariância entre estes efeitos, tanto sob metodologia frequentista quanto *bayesiana*. Todavia, este tipo de abordagem introduz considerável complexidade nos algoritmos para estimação de componentes de variância, além de não ser possível sua implementação nos pacotes usualmente empregados para avaliação genética de animais domésticos, além do que alguns conjuntos de dados não permitiriam a sua utilização devido a problemas de superparametrização/estimabilidade.

Frente a estas dificuldades, desenvolveu-se uma abordagem em dois passos, como meio de investigar diferenças de ordem genética entre touros para a variância residual de sua progênie, fazendo uso dos resíduos preditos para as observações de ganho de peso do nascimento à desmama de sua progênie, após a remoção de efeitos genéticos e ambientais sobre a média desta característica.

### 3.2.MATERIAL E MÉTODOS

Optou-se por utilizar os dados da base Aliança Nelore, pelo fato deste conjunto de dados apresentar maior número de observações e maior informação de parentesco disponível (325.432 animais, 3.339 touros, 985 pais de touros e 330 mães de touros), comparado ao conjunto de dados do programa PAINT (utilizado em estudos preliminares deste trabalho e que também faz parte da base de dados Aliança).

A característica ganho de peso do nascimento à desmama (GND, em kg) foi escolhida para este estudo sobre heterogeneidade genética na variância residual de características de interesse econômico em bovinos da raça Nelore. GND é um componente do peso em idades posteriores, sendo que um bezerro desmamado, pesando entre 190 e 200 kg, também constitui a unidade de comercialização do mercado de cria na maioria das regiões do país, implicando na existência de um intervalo ótimo para esta característica. Sendo assim, a relevância da uniformidade de produção para esta característica e o fato de ser um critério de seleção bastante estudado na população em questão, constituindo o maior banco de dados disponível, foram preponderantes para a utilização de GND neste estudo.

Dados de 325.432 animais com observações de GND e cujos pais fossem conhecidos foram utilizados no primeiro passo deste estudo. Tais dados foram provenientes de 385 diferentes rebanhos distribuídos no Brasil Central e foram coletados entre 1983 e 2007.

No primeiro passo, os dados foram analisados segundo procedimentos convencionais para estimação de variância aditiva sobre a média de GND, após terem sido pré-ajustados para os efeitos fixos de idade do animal (padronização para idade de 205 dias), idade da vaca ao parto (padronização para idade de 7 anos) e data juliana média dentro da estação de nascimento, sendo as observações resultantes referidas como ganho do nascimento à desmama pré-ajustado (GNDa). Foram utilizados fatores de ajuste previamente estimados com dados da base Aliança Nelore.

Estimativas para componentes de variância e soluções para os efeitos fixos e aleatórios sobre GNDa foram obtidas através do método da máxima verossimilhança restrita (REML), utilizando-se o do *software* MTDFREML (BOLDMAN et al., 1995), em que se implementou o seguinte modelo animal:

$$GNDa = Xgc + Z1ad + Z2am + Z3ap + e, \text{ (Modelo 1)}$$

considerando o efeito fixo de grupo de contemporâneos (gc, determinado pela concatenação das informações de rebanho, ano, estação, sexo, data de pesagem e grupo de manejo) e os efeitos aleatórios aditivo direto (**ad**), aditivo materno (**am**), de ambiente permanente devidos à vaca (**ap**) e residual (**e**), sendo **X**, **Z1**, **Z2** e **Z3** as respectivas matrizes de incidência de cada efeito sobre GNDa.

Caracterizando as pressuposições deste modelo, as matrizes de esperança e (co)variância dos efeitos considerados são dadas, respectivamente, por:

$$E(\mathbf{GND}) = Xgc; E(\mathbf{ad}) = E(\mathbf{am}) = E(\mathbf{ap}) = E(\mathbf{e}) = 0$$

$$VAR \begin{pmatrix} \mathbf{ad} \\ \mathbf{am} \\ \mathbf{ap} \\ \mathbf{e} \end{pmatrix} = \begin{pmatrix} A\sigma_{ad}^2 & A\sigma_{dm} & 0 & 0 \\ A\sigma_{dm} & A\sigma_{am}^2 & 0 & 0 \\ 0 & 0 & I\sigma_{ap}^2 & 0 \\ 0 & 0 & 0 & I\sigma_e^2 \end{pmatrix}.$$

Onde  $\sigma_{ad}^2$  é a variância genética aditiva para efeito direto,  $\sigma_{am}^2$  é a variância genética aditiva para efeito materno,  $\sigma_{dm}$  é a covariância entre os efeitos aditivos direto e materno,  $\sigma_{ap}^2$  é a variância do efeito de ambiente permanente devido à mãe e  $\sigma_e^2$  é a variância residual, **A** é a matriz de parentesco entre os indivíduos e **I** é uma matriz identidade de ordem igual ao número de observações.

Como 84,77% das vacas com informação para desempenho de progênie nesta base de dados não possuíam informação de desempenho próprio para a característica em estudo, optou-se por também assumir covariância nula entre efeito aditivo direto e aditivo materno sobre GNDa ( $\sigma_{dm} = 0$ ), uma vez que, sob esta estrutura de *pedigree*, seriam esperadas estimativas bastante viesadas desta covariância, conforme apontado no estudo de simulação conduzido por HEYDARPOUR et al. (2008). Os demais componentes de covariância entre os efeitos aleatórios foram assumidos iguais a zero.

O segundo passo deste estudo iniciou-se com o uso das soluções para efeitos fixos e aleatórios obtidas sob o Modelo 1 para obter os resíduos preditos ( $\hat{e}$ ) associados a cada observação de GNDa.

Para obterem-se estimativas de dispersão residual dentro de famílias de touro, calcularam-se seis medidas de dispersão dos resíduos da progênie de cada touro: desvio-padrão (SD), variância (VAR), log de VAR (ln), *median absolute deviation* (MAD), *Qn scale* (Qn) e *Sn scale* (Sn), sendo que as três últimas são consideradas medidas robustas de dispersão e poderiam apresentar propriedades interessantes para tratar o problema em questão (ROUSSEEUW & CROUX, 1993), sobretudo no caso de touros com progênie de pequeno tamanho e/ou com filhos pertencentes a grupos de contemporâneos mais variáveis. Todas as medidas de dispersão residual foram obtidas por meio do procedimento UNIVARIATE do pacote estatístico SAS, usando-se a opção *robustscale* (SAS INSTITUTE, 2002).

Com o objetivo de investigar a influência do tamanho da progênie sobre as estimativas de variância aditiva para dispersão residual, consideraram-se dois conjuntos de dados no segundo passo das análises:

A) conjunto contendo dados de dispersão da progênie de 985 touros com, no mínimo, 10 filhos com dados de GND.

B) conjunto contendo dados de dispersão da progênie de 758 touros com, no mínimo, 25 filhos com dados de GND.

Todos os touros considerados no segundo passo possuíam pai conhecido e havia informação de avô paterno para 60,41% e 60,69% deles, respectivamente, nos conjuntos A e B.

Para cada conjunto de dados e estimador de dispersão residual, o seguinte modelo foi utilizado para obter estimativas REML da variância aditiva sobre a dispersão residual na escala de cada estimador, usando o *software* MTDFREML:

$$d = Z_a g + \varepsilon \quad (\text{Modelo 2}),$$

em que  $\mathbf{d}$  é o vetor de medidas de dispersão residual da progênie dos touros,  $\boldsymbol{\mu}$  é a dispersão residual média,  $\mathbf{g}$  é o vetor de efeito aditivo de pai de touro e  $\boldsymbol{\varepsilon}$  é o vetor de efeitos residuais (no segundo passo) e  $\mathbf{Z}_a$  é a matriz de incidência que relaciona os elementos de  $\mathbf{g}$  a  $\mathbf{d}$ .

As matrizes de esperança e (co)variância são dadas por:

$$E(d) = \boldsymbol{\mu} \quad (\text{i.e. dispersão residual média}), \quad E(g) = E(\varepsilon) = 0; \quad \text{Variância } (d) = A\sigma_g^2 + I\sigma_\varepsilon^2,$$

onde  $A$  é a matriz de parentesco entre os pais de touros (baseada nos avôs paternos dos

touros) e  $I$  é uma matriz identidade. Neste modelo assume-se  $\mathbf{g}$  e  $\boldsymbol{\varepsilon}$  normalmente distribuídos, i.e.  $\mathbf{g} \sim N(0, A\sigma_g^2)$  e  $\boldsymbol{\varepsilon} \sim N(0, I\sigma_\varepsilon^2)$ , e que  $\sigma_g^2$  equivale **1/16** da variância aditiva para dispersão residual ( $\sigma_{Av}^2$ ), uma vez que as observações de resíduo obtidas no primeiro passo das análises referem-se aos netos dos animais aos quais o efeito  $\mathbf{g}$  é associado.

Adotou-se a variância da função *simplex* menor do que  $10^{-12}$  como critério de convergência. Inicialmente, neste segundo passo, considerou-se diretamente o efeito de touro ao invés do efeito de pai de touro no Modelo 2, porém devido a problemas de convergência, optou-se por não usar este tipo de modelo, pois ao se considerar o efeito de touro, as estimativas mostraram-se sensíveis ao uso de diferentes *priors*. Uma possível explicação para tal ocorrência seria que, mesmo com um conjunto de dados bastante grande no primeiro passo, não foi possível recuperar informações de parentesco suficiente entre touros para garantir estimativas confiáveis dos componentes de variância sob este modelo.

Como meio de obter estimativas dos componentes de variância no segundo passo sob um procedimento de estimação alternativo, o Modelo 2 também foi implementado por meio da macro GLIMMIX no *software* SAS (LITTELL et al., 1996), a fim de realizar uma análise de componentes de variância bayesiana não conjugada (NBVC). Um modelo normal linear (sem considerar a matriz de parentesco) foi usado para simular 10.000 amostras pseudo-aleatórias dos componentes de variância, permitindo que a média *a posteriori* da variância aditiva fosse estimada com desvio-padrão equivalente a 1/100 do desvio padrão deste parâmetro.

A abordagem de NBVC faz uso de *priors* não informativos (*Jeffrey's prior*) e um algoritmo de cadeia independente, sendo que, deste modo, não é necessário a determinação de período de *burn-in* (WOLFINGER & KASS, 2000). Utilizou-se o procedimento KDE do SAS (SAS Institute, 2002) para estimar a densidade posterior das variâncias aditivas das diferentes medidas de dispersão dos resíduos preditos da progênie dos touros, conforme também utilizado por GUERRA et al. (2006).

Como os estimadores de dispersão usados neste estudo possuem diferentes escalas, assim como, suas respectivas estimativas de variância aditiva, optou-se por transformar as estimativas obtidas para escala da variância (i.e. a mesma escala de VAR), a fim de facilitar as comparações entre estimadores, assim como para facilitar as

inferências acerca da oportunidade de resposta à seleção para incremento na uniformidade de GNDa.

A fórmula proposta por MULDER et al. (2007) foi utilizada para converter as estimativas de  $\sigma^2_{Av}$  obtidas com os dados de ln para a escala de VAR. No caso dos estimadores de escala, a transformação foi realizada multiplicando-se cada estimativa de variância aditiva sobre a dispersão residual por um fator  $f$ , calculado como a razão entre as variâncias de VAR e SD, fazendo-se uso de expressões derivadas por LYNCH & WALSH (1998) e assumindo-se que os resíduos associados a cada observação de GNDa apresentavam distribuição normal (Apêndice I).

Para cada estimador, calculou-se uma medida da variabilidade genética para dispersão residual seguindo conceito similar à *evolvability* (Ev) (HOULE, 1992):

$$Ev = \sigma_{Av} / \mu,$$

em que  $\sigma_{Av}$  é o desvio padrão aditivo para dispersão residual e  $\mu$  é a dispersão residual média. As estimativas de Ev foram obtidas tanto na escala original de cada estimador quanto transformadas para a escala de VAR.

Calculou-se a herdabilidade da variância residual ( $h^2_v$ ), conforme proposto por MULDER et al. (2007), a fim de obter-se um parâmetro para estimativa da acurácia da seleção com base em Av. Sendo assim, para cada estimador,  $h^2_v$  foi calculada como:

$$h^2_v = \sigma^2_{Av} / (2 \sigma^4_P + 3 \sigma^2_{Av}),$$

em que  $\sigma^4_P$  é o quadrado da variância fenotípica e  $\sigma^2_{Av}$  é a variância aditiva para dispersão residual (transformada para a escala da variância).

Como a correlação entre Am e Av não é estimável diretamente sob a abordagem proposta no presente estudo, a correlação entre as DEPs dos touros para dispersão residual e para a média de GNDa poderia ser usada como um indicativo da correlação genética entre a média e a dispersão residual, mas que dependeria das acurácias destes dois tipos de DEP.

Sendo assim, foi empregado o procedimento proposto por CALO et al. (1973) para ajustar a estimativa desse tipo de correlação com base nas acurácias de predição das DEPs, o qual também foi empregado por WOLC et al. (2009) para estudar a correlação genética entre média e variância residual do peso vivo de frangos de corte. No presente estudo, as acurácias de predição das DEPs para média de GNDa foram obtidas por meio

do software MTDFREML. A acurácia das DEPs de dispersão foi calculada por meio da fórmula proposta por MULDER et al. (2007), que leva em conta a  $h^2_v$  obtida para cada estimador e o tamanho da progênie de cada touro.

Como uma alternativa ao uso de transformações da variância aditiva no segundo passo para a escala da variância, a fim de comparar os estimadores quanto à oportunidade de resposta à seleção (*evolvability*), calculou-se a redução que seria esperada na variância residual de GNDa, com base na redução na dispersão residual predita na escala de cada estimador, após uma geração de seleção, com base na fórmula proposta por BIJMA (2009):

$$\frac{\Delta\mu}{\mu} = i * r_{\hat{A}_v, A_v} * Ev$$

em que  $\mu$  é a dispersão residual média antes da seleção,  $\Delta\mu$  é a mudança esperada em  $\mu$ ,  $i$  é a intensidade de seleção,  $r_{\hat{A}_v, A_v}$  é a acurácia média de predição de  $A_v$  e  $Ev$  é a *evolvability* para dispersão residual.

Assumindo mesma intensidade de seleção e acurácia de predição para todos os estimadores, a redução na dispersão residual na escala de cada estimador foi calculada e, com base nesta, a redução esperada na variância residual, tornando possível uma comparação empírica entre valores de *evolvability* em diferentes escalas.

### 3.3. RESULTADOS e DISCUSSÃO

#### *Estatísticas descritivas referentes às observações de GND e de sua dispersão residual*

Para o conjunto de dados utilizados neste estudo, a média de GND foi igual a 141,58 kg, com coeficiente de variação igual a 18,83%. Após o pré-ajuste para os efeitos fixos descritos anteriormente, os dados apresentaram média 170,32 kg, com coeficiente de variação igual a 23,17%.

Os valores de assimetria e curtose da distribuição de GNDa foram, 0,152 e -0,247, respectivamente, fornecendo indicativo de uma distribuição aproximadamente normal dos dados utilizados para a estimação de componentes de variância no primeiro passo do estudo.

As estatísticas descritivas das medidas de dispersão residual da progênie dos touros (i.e. os dados efetivamente utilizados no segundo passo das análises) são apresentadas na Tabela 7, para os dois conjuntos de dados e seis estimadores de dispersão residual considerados.

A imposição de restrição quanto ao tamanho mínimo de progênie para consideração de um touro no segundo passo deste estudo implicou em tamanhos médios de progênie iguais a 136 e 193, respectivamente, para os conjuntos de dados A e B.

O aumento no tamanho mínimo de progênie levou a uma redução no coeficiente de variação de todos os estimadores, porém teve pouco impacto sobre a dispersão residual média, indicando que considerar touros com, no mínimo 25 filhos, implicou na obtenção de um conjunto de dados com menor variabilidade para dispersão residual, sendo este efeito mais pronunciado para os estimadores robustos e ln.

Para a maioria dos estimadores de dispersão estudados, os dados não apresentaram distribuição normal, com marcado desvio da normalidade para VAR (Tabela 8). De modo geral, as distribuições no conjunto de dados B foram mais próximas da normalidade, o que poderia refletir uma menor influência de valores discrepantes, ao se fixar um maior tamanho de progênie para a inclusão de observações neste conjunto. Deste modo, espera-se que as estimativas de componentes de variância no conjunto B sejam mais informativas, tanto pelo menor desvio da normalidade, quanto pelo menor viés no cálculo das medidas de dispersão da progênie de cada touro.

Tabela 7 - Estatísticas descritivas da dispersão residual da progênie dos touros para GNDa, considerando dois tamanhos mínimos de progênie e seis estimadores de dispersão

Conjunto de dados	A		B	
Tamanho mínimo de progênie	10		25	
Estimador de dispersão residual	Média	CV	Média	CV
Desvio-padrão (SD)	14,06	0,243	14,18	0,205
Variância (VAR)	209,36	0,502	209,49	0,423
Log de VAR (ln)	5,23	0,096	5,26	0,078
<i>Median absolute deviation</i> (MAD)	8,41	0,284	8,47	0,230
<i>Qn scale</i> (Qn)	14,45	0,278	13,82	0,207
<i>Sn scale</i> (Sn)	12,91	0,265	12,95	0,212

CV=coeficiente de variação; GNDa = ganho de peso do nascimento à desmama pré-ajustado, em kg.

Tabela 8 - Estatísticas referentes à assimetria e à curtose das distribuições dos dados de dispersão da progênie dos touros para GND, considerando dois conjuntos de dados e seis estimadores de dispersão

Conj. de dados	A		B	
	Estimador	Assimetria	Curtose	Assimetria
VAR	1,4046	3,9085	1,3483	4,2598
SD	0,4867	1,1162	0,5548	1,0150
ln	-0,5472	2,2199	-0,0724	0,0938
MAD	0,6805	2,4284	0,4250	0,8828
Qn	1,4050	5,0718	0,6517	1,0448
Sn	0,6363	2,3665	0,4343	0,7086

GNDa = ganho de peso do nascimento à desmama pré-ajustado, em kg; A = mínimo de 10 filhos por touro; B = mínimo de 25 filhos por touro

A rigor, apenas a distribuição de ln no conjunto B foi considerada normal (com base nos testes de Shapiro-Wilk e Kolmogorov-Smirnov), todavia as distribuições de SD, MAD e Sn neste conjunto também foram próximas da normalidade, apesar dos valores positivos para assimetria e distribuição leptocúrtica.

#### *Parâmetros genéticos para a média de GNDa*

Após as análises realizadas no primeiro passo do estudo (considerando o Modelo 1), foram estimadas variâncias de 105,85 kg<sup>2</sup> e 24,83 kg<sup>2</sup> para os efeitos genéticos aditivos direto e materno, respectivamente. As estimativas de herdabilidade foram iguais a 0,20±0,007 para o efeito aditivo direto e 0,05±0,005 para o efeito aditivo materno.

A variância do efeito de ambiente permanente foi igual a 80,38 kg<sup>2</sup>. A variância residual foi estimada em 311,27 kg<sup>2</sup>, o que corresponde a 59% da variabilidade fenotípica de GNDa e foi superior a variância residual efetivamente considerada no segundo passo das análises (aproximadamente 209 kg<sup>2</sup>), em virtude da restrição quanto ao tamanho mínimo de progênie dos touros.

De modo geral, os resultados obtidos no primeiro passo das análises, referentes aos parâmetros associados à média de GNDa, estão em concordância com os de estudos anteriores que obtiveram estimativas de herdabilidade de magnitude baixa a moderada para esta característica (e.g. MARCONDES et al. 2000; ORTIZ PEÑA et al., 2004). Tal resultado também indicaria oportunidade de resposta à seleção para incremento na média de GND.

*Evidências de efeitos aditivos sobre a dispersão residual de GNDa*

Para cada estimador de dispersão residual, as estimativas de variância aditiva obtidas no segundo passo do estudo são apresentadas na tabela 9, separadas por conjunto de dados e procedimento de estimação empregado.

A despeito das diferentes escalas entre os estimadores de dispersão, houve razoável evidência de variância aditiva não nula para todos os estimadores (o valor 0 não foi compreendido pela região de credibilidade de  $\sigma^2_{Av}$  em todas as situações).

Para um mesmo conjunto de dados (i.e. mesmo tamanho mínimo de progênie), as estimativas de  $\sigma^2_{Av}$  foram consistentes tanto por REML quanto por inferência bayesiana, com estimativas pontuais não diferindo em mais de 6% para um mesmo estimador (exceto no caso de MAD, cuja estimativa REML de variância aditiva foi 13% superior no conjunto de dados B).

Tabela 9 - Estimativas de variância aditiva para dispersão residual de GNDa, obtidas por máxima verossimilhança restrita (REML) e análise de componentes de variância bayesiana não conjugada (NBVC), para dois conjuntos de dados e seis diferentes estimadores de dispersão, sob o Modelo 2

Procedimento	REML		NBVC					
	Conjunto de dados		A			B		
Estimador	A	B	$\sigma^2_{Av}$	low	up	$\sigma^2_{Av}$	low	up
SD	15,6	14,28	15,68	6,24	25,92	14,48	5,84	23,76
VAR	10908	9841	10344	3220	17724	9576	3524	16356
In	0,44	0,36	0,44	0,196	0,72	0,36	0,168	0,6
MAD	7,84	5,56	7,8	2,96	12,92	4,92	1,16	8,76
Qn	22,52	16,08	22,52	11,2	34,88	16	7,2	25,52
Sn	17,92	12	18,84	8,6	29,4	12,04	4,08	20,32

GND = ganho de peso do nascimento à desmama pré-ajustado, em kg. A = mínimo de 10 filhos por touro; B = mínimo de 25 filhos por touro;  $\sigma^2_{Av}$  = estimativa de variância aditiva para dispersão residual; low e up = limite inferior e superior para região de credibilidade com 95% de probabilidade para  $\sigma^2_{Av}$ , respectivamente.

O aumento no tamanho mínimo de progênie implicou em estimativas inferiores de variância aditiva da dispersão residual, sobretudo no caso de estimadores robustos, cujas estimativas de  $\sigma^2_{Av}$  foram cerca de 30% menores. Tal resultado poderia ser associado tanto a um menor viés nas estimativas de dispersão por touro, por considerarem-se famílias de maior tamanho, quanto a uma redução sistemática na variabilidade, uma vez que menos famílias foram consideradas no segundo passo.

As estimativas de *evolubility* (Ev), calculadas com estimativas de  $\sigma^2_{AV}$  transformadas para a escala de variância, são apresentadas na Tabela 10, agrupadas por estimador e conjunto de dados.

Como as estimativas de dispersão residual média praticamente não diferiram entre os conjuntos de dados A e B (Tabela 7), as menores estimativas de variância aditiva obtidas considerando-se famílias com, no mínimo, 25 produtos (conjunto B) implicaram em menores estimativas de *evolubility* para todos os estimadores, sendo que as maiores reduções foram observadas no caso de MAD e Sn.

As estimativas de *evolubility* calculadas para um mesmo estimador e conjunto de dados com base em estimativas de  $\sigma^2_{AV}$  por REML e NBVC também apresentaram valores bastante próximos, indicando que a consideração da informação de parentesco entre touros e/ou a metodologia de estimação pouco afetaram as estimativas no segundo passo.

Tabela 10 - *Evolubility* (Ev) para dispersão residual de GNDa, calculada com estimativas de variância obtidas por máxima verossimilhança restrita (REML) e análise de componentes de variância bayesiana não conjugada (NBVC), para dois conjuntos de dados e seis diferentes estimadores de dispersão

Procedimento	REML		NBVC	
	A	B	A	B
	Ev	Ev	Ev	Ev
SD	56,18%	53,30%	56,33%	53,67%
VAR	49,89%	47,35%	48,58%	46,71%
In	73,75%	66,00%	73,80%	66,00%
MAD	44,91%	37,55%	44,80%	35,33%
Qn	65,68%	58,03%	65,68%	57,89%
Sn	65,58%	53,50%	67,24%	53,59%

GND = ganho de peso do nascimento à desmama pré-ajustado, em kg. A = mínimo de 10 filhos por touro; B = mínimo de 25 filhos por touro

Calculando-se a *evolubility* com as estimativas de variância aditiva e dispersão média na escala original de cada estimador foi possível notar uma marcante diferença de escala entre os estimadores, o que justificaria o uso da transformação para comparação entre estimadores. Em sua escala original, as estimativas de Ev oscilaram entre 46 e 50% no caso de VAR, 11 e 12% no caso de In e foram próximas de 30% para os demais estimadores.

As estimativas de *evolability* após a transformação da variância aditiva para a escala de VAR foram bastante elevadas (acima de 40% na maioria dos casos). Os maiores valores foram obtidos para In, ao passo que para o estimador robusto MAD calcularam-se os mais baixos valores para este parâmetro. Em estudo de revisão, MULDER et al. (2007) indicaram que a maioria das estimativas de *evolability* para a variância residual de características de produção em diversas espécies situou-se entre 25% e 60%.

Os maiores valores de *evolability* encontrados no presente estudo podem se dever a uma superestimação da variância aditiva sobre a dispersão residual no segundo passo, pelo fato desta abordagem não contemplar a influência de efeitos ambientais sobre a dispersão da progênie dos touros (e.g. heterogeneidade de variância residual entre grupos de contemporâneos), além do que apenas uma pequena fração da variância aditiva total (1/16) pode ser explicada pelo efeito aditivo considerado no Modelo 2. Do mesmo modo, os próprios procedimentos de transformação podem ter constituído uma fonte adicional de viés, a despeito de permitirem comparação entre estimadores em diferentes escalas.

Na tabela 11, apresentam-se os valores de herdabilidade da variância residual ( $h^2_v$ ) calculados com base nas estimativas REML de  $\sigma^2_{Av}$  obtidas no segundo passo das análises, para cada estimador e conjunto de dados. Na mesma tabela, apresenta-se a acurácia média de predição do valor genético dos touros para variância residual ( $r_{AV,\hat{AV}}$ ), considerando o tamanho médio de progênie em cada conjunto de dados e o respectivo valor de  $h^2_v$ .

Para ambos os conjuntos de dados, os valores de  $h^2_v$  foram baixos para todos os estimadores (inferiores a 0,04), ao passo que o tamanho médio de progênie dos touros pode ser considerado alto, sendo igual a 136 (restrição de, no mínimo, 10 filhos por touro) e a 193 (restrição de, no mínimo, 25 filhos por touro) nos dois conjuntos de dados de dispersão da progênie considerados no estudo. Sendo assim, as acurácias médias de predição apresentaram valores expressivos em ambos os conjuntos de dados para a maioria dos estimadores (valores maiores que 0,6, exceto no caso de MAD).

Tais resultados indicariam que, apesar dos baixos valores de  $h^2_v$  que foram estimados, a consideração de dados de dispersão da progênie de animais com número elevado de filhos (e.g. 100) poderia permitir predições acuradas de seu mérito para

variabilidade da progênie, como também indicado nos estudos de MULDER et al. (2009) e WOLC et al. (2009) com dados frangos de corte.

Os maiores valores de  $h^2v$  foram observados no caso de In e Qn, em ambos os conjuntos de dados. Uma vez que a acurácia de predição depende de  $h^2v$  e do tamanho da progênie dos touros, espera-se que esses estimadores forneçam estimativas mais acuradas de  $Av$ .

Tabela 11 – Estimativas de herdabilidade da variância residual de GNDa ( $h^2v$ ) e acurácia média de predição de  $Av$  dos touros ( $r_{\hat{A}v, Av}$ ), considerando dois tamanhos mínimos de progênie

Conjunto de dados	A (n=136)		B (n=193)		
	Estimador	$h^2v$	$r_{\hat{A}v, Av}$	$h^2v$	$r_{\hat{A}v, Av}$
	SD	0,0212	0,6479	0,0198	0,7118
	VAR	0,0189	0,6260	0,0171	0,6911
	In	0,0386	0,7550	0,0317	0,8081
	MAD	0,0087	0,4782	0,0063	0,5442
	Qn	0,0312	0,7190	0,0211	0,7766
	Sn	0,0205	0,6423	0,0141	0,7065

GNDa = ganho de peso do nascimento à desmama pré-ajustado, em kg.

A =mínimo de 10 filhos por touro; B = mínimo de 25 filhos por touro

$h^2v = \sigma^2_{Av} / (2 \sigma^2_p + 3 \sigma^2_{Av})$ , onde  $\sigma^2_{Av}$  é a variância aditiva para variância residual.

$r_{\hat{A}v, Av} = (1/4 * n * h^2v / (1 + 1/4 (n - 1) * h^2v))^{1/2}$ , onde  $n$  é o tamanho médio de progênie.

Apesar dos menores valores obtidos para  $h^2v$  no conjunto de dados B (em virtude das menores estimativas obtidas para  $\sigma^2_{Av}$  com o aumento do tamanho mínimo de progênie), as acurácias de predição seriam, em média, superiores, em virtude do maior número médio de filhos por touro neste conjunto de dados.

Na tabela 12, apresenta-se a redução esperada na variância residual após uma geração de seleção, calculada segundo a equação proposta por BIJMA (2009), fazendo uso dos valores de *evolability* para dispersão residual na escala original de cada estimador e das acurácias médias de predição (Tabela 11). Nestes cálculos, assumiu-se seleção de 20% dos touros candidatos com base no seu mérito genético para dispersão residual, bem como ausência de seleção para dispersão residual entre as vacas.

É importante ressaltar que estes resultados devem ser considerados otimistas, tendo em vista que se está contemplando seleção exclusivamente para redução na dispersão residual. Na prática, a intensidade de seleção seria bem menor, considerando seleção

concomitante para outros critérios, sobretudo se estes forem correlacionados de maneira desfavorável com a dispersão residual. Além disso, conforme discutido anteriormente, as estimativas de *evolubility* podem ter sido superestimadas pelo fato de não ter sido possível considerar efeitos ambientais no segundo passo das análises. Sendo assim, tais resultados serviriam principalmente como uma fonte adicional de comparação entre os diferentes estimadores estudados.

Sob estas circunstâncias, considerável resposta para redução na variância residual de GND seria esperada (entre 20% e 30% por geração), sobretudo no caso de In, Qn e Sn. Praticamente não houve diferenças entre a predição de resposta a seleção nos dois conjuntos de dados, o que poderia ser explicado pelo fato da maior acurácia de predição obtida com maior tamanho de progênie no conjunto B ter sido contrabalanceada pelas menores estimativas de *evolubility* obtidas neste caso.

Tabela 12 - Redução esperada\* na variância residual de GNDa após uma geração de seleção, para os dois conjuntos de dados de dispersão residual da progênie de touros da raça Nelore

Estimador	Conjunto de dados	
	A	B
SD	23,86%	24,79%
VAR	21,86%	22,91%
In	29,57%	28,78%
MAD	21,05%	20,08%
Qn	30,32%	29,05%
Sn	27,31%	24,70%

\*Assumiu-se seleção de 20% dos touros candidatos com base no seu mérito genético para dispersão residual e ausência de seleção para dispersão residual entre as vacas. A = mínimo de 10 filhos por touro; B = mínimo de 25 filhos por touro

As estimativas de correlação entre DEPs para efeito direto sobre a média de GNDa e para dispersão residual de GND são apresentadas na Tabela 13, assim como as correlações ajustadas pela acurácia de cada tipo de DEP.

Nenhum estimador de dispersão residual apresentou DEPs com correlação significativa com DEPs para efeito materno, motivo pelo qual as estimativas ajustadas pela acurácia das DEPs não foram apresentadas neste caso.

Correlações negativas, de magnitude baixa, porém significativas ( $P < 0,0001$ ) foram observadas entre DEPs para média e as DEPs de todos estimadores de dispersão residual.

Via de regra, após o ajuste com base na acurácia de predição das diferentes DEPs, o valor absoluto da correlação aumentou ligeiramente (entre 20 e 28%), mantendo o mesmo sinal.

Além disso, as correlações entre DEPs para efeito direto e de dispersão foram maiores ao se considerar o conjunto de dados B (maior tamanho de progênie), sobretudo no caso de MAD e Sn, porém tais correlações ainda podem ser consideradas baixas.

Tabela 13 - Correlações entre DEPs para a média (efeito direto e materno) e as DEPs para dispersão residual de GNDa, para dois conjuntos de dados de dispersão da progênie e usando seis diferentes estimadores de dispersão

Conjunto de dados A						
DEPs para média	DEPs para dispersão residual					
	VAR	SD	In	MAD	Qn	Sn
Efeito direto	-0,1724*	-0,1649*	-0,1515*	-0,1244*	-0,1886*	-0,1420*
Efeito direto**	-0,2155	-0,2116	-0,1847	-0,1541	-0,2335	-0,1745
Efeito materno	0,0017	0,0048	0,0095	0,0172	-0,0043	0,0175
Conjunto de dados B						
DEPs para média	DEPs para dispersão residual					
	VAR	SD	In	MAD	Qn	Sn
Efeito direto	-0,1818*	-0,1759*	-0,1671*	-0,1771*	-0,1999*	-0,1899*
Efeito direto**	-0,2274	-0,2258	-0,2038	-0,2193	-0,2475	-0,2334
Efeito materno	-0,0156	-0,0138	-0,0104	-0,0014	-0,0206	-0,0099

A = mínimo de 10 filhos por touro; B = mínimo de 25 filhos por touro

\*p-value < 0.0001 \*\*Estimativas ajustadas pela acurácia das DEPs, segundo procedimento descrito por Calo et al. (1973).

WOLC et al. (2009) aplicaram procedimento análogo ao do presente estudo para estimar a correlação genética entre a média e a variância residual do peso corporal de frangos de corte. Estes autores também encontraram associação negativa e de baixa magnitude entre a média e a variância residual. Estimativas similares de correlação genética também foram obtidas por meio de procedimentos diretos de estimação para características de peso em frangos de corte (ROWE et al., 2006; MULDER et al., 2009).

Na literatura, não há evidências de um padrão geral para a associação entre efeitos genéticos sobre a média e a variância residual de características ligadas ao peso corporal de animais domésticos. Em alguns estudos, estimaram-se correlações elevadas e positivas entre estes efeitos (ROS et al., 2004; GUTIERREZ et al. 2006), ao passo que estimativas negativas de magnitude igualmente elevada também foram relatadas por outros autores (e.g. HUBY et al., 2003; GUTIERREZ et al. 2006), indicando a necessidade de maiores

estudos para investigar a influência de efeitos genéticos, ambientais, além de efeitos de escala sobre este tipo de parâmetro.

A despeito da baixa magnitude, a associação negativa entre as DEPs para média e para dispersão residual de GNDa encontrada no presente estudo seria favorável em situações em que uma maior uniformidade é desejada e indicaria a possibilidade de se aumentar a média de GND e reduzir sua variabilidade simultaneamente, por meio de seleção.

A resposta à seleção baseada em índices combinando o mérito genético para a média e a variância residual de uma característica simulada foram estudados por MULDER et al. (2008). Segundo estes autores, a inclusão da variância residual no objetivo de seleção seria relevante para características com equações de lucro não lineares, mas via de regra, maiores reduções na variância residual (i.e. incremento na uniformidade) seriam obtidas quando a média da população se encontrasse próxima do ótimo.

As estimativas de variância aditiva para dispersão residual no presente estudo constituíram uma evidência empírica de heterogeneidade genética sobre a variância residual de GNDa, como também relatado para outras características ligadas a peso corporal em animais de produção (e.g. IBAÑEZ-ESCRICHE et al., 2008; WOLC et al, 2009). Além disso, a considerável concordância entre estimativas comparáveis obtidas por REML e inferência bayesiana neste estudo pode ser considerada um indício de que a variância aditiva estimada não seria apenas um artefato numérico decorrente do processo de estimação.

A abordagem em dois passos empregada neste estudo permitiria o estudo da variabilidade genética sobre a variância residual em bovinos de corte sem grande complexidade do ponto de vista computacional. Todavia, muitas pressuposições devem ser atendidas para garantir a confiabilidade dos resultados obtidos (e.g. normalidade das medidas de dispersão residual, ausência de efeitos de heterogeneidade devido a GC).

Dado o pressuposto de normalidade no segundo passo das análises, deve-se considerar que os resultados obtidos para os estimadores com distribuições mais próximas de uma normal (i.e. In, SD e MAD) sejam mais informativos.

Do mesmo modo, dada a maior proximidade da normalidade para os dados provenientes da dispersão de famílias de touro com, no mínimo, 25 filhos, pode-se esperar

que a consideração de famílias com este tamanho mínimo seja mais adequada para implementação da presente abordagem, além do fato de que estimativas de dispersão obtidas com maior número de filhos sejam mais confiáveis.

Como alguns efeitos ambientais estimáveis influenciam a média de GND (e.g. rebanho, sexo, data juliana dentro da estação de nascimento) é razoável esperar que tais efeitos também possam influenciar em alguma medida a variância residual desta característica. Além disso, alguns estudos sobre heterogeneidade genética sobre a variância relataram melhor ajuste para modelos considerando a ação de efeitos fixos sobre a variância residual concomitantemente com efeitos aditivos (SORENSEN & WAAGEPETERSEN, 2003; IBAÑEZ-ESCRICHE, 2008).

A existência de heterogeneidade entre GCs para variância residual não pode ser investigada pelo Modelo 2 deste estudo, uma vez que sob esta abordagem não há informação suficiente para decompor as medidas de dispersão residual entre efeitos de GC e de touro, pois o pequeno número médio de filhos por touro num mesmo GC inviabilizaria calcularem-se medidas de dispersão residual na progênie dentro de cada GC.

MULDER et al. (2008) argumentaram que, em presença de heterocedasticidade devida a efeitos ambientais, seriam necessárias famílias de tamanho ainda maior do que o necessário para estimar o mérito genético para dispersão quando se pressupõe a ação exclusiva de efeitos aditivos sobre a variância residual.

As limitações identificadas na presente abordagem reiteram a necessidade de mais estudos sobre este tema, por exemplo, usando-se distribuições alternativas para o erro dos dados de dispersão da progênie sob modelos lineares mistos generalizados (GLMM) ou mesmo, empregando-se modelos alternativos que permitissem contemplar a influência de efeitos fixos no segundo passo das análises, tais como aqueles empregados por WOLC et al. (2009) e MULDER et al. (2009) para dados de frangos de corte.

No presente estudo, quase 60% da variabilidade fenotípica de GNDa deveu-se a efeitos ligados à variância residual. Assumindo-se que a ação aditiva sobre a variância residual desta e de outras características ligadas ao peso de bovinos de corte apresente magnitude similar ao que tem sido descrito na literatura, pode-se esperar que a seleção de touros para menor variância residual, com base em dados de progênie de tamanho

razoável (digamos 100 filhos) possa permitir incrementos significativos na uniformidade destas características.

*Estudo sobre possíveis efeitos de prepotência dos touros sobre a variabilidade da progênie*

Após a realização do primeiro passo das análises, investigou-se um procedimento alternativo para definição dos resíduos preditos usados para estimar parâmetros genéticos da dispersão residual, como meio de avaliar uma eventual influência de efeitos de prepotência dos touros sobre a variabilidade da progênie (i.e. formação de gametas menos variáveis por animais endogâmicos), apesar deste efeito ser pouco provável para características com herança poligênica (HOHENBOKEN, 1985; BIJMA, 2009).

Neste caso, os resíduos (RESm) de cada observação foram calculados com base no efeito de segregação mendeliana devido a touro somado ao efeito residual. As estimativas de segregação mendeliana devida a touro (SMt) foram obtidas por  $SMt = \frac{1}{2} * [EBVp - ((EBVt + EBVv)/2)]$ , em que EBVp, EBVt e EBVv são os valores genéticos preditos de cada produto, touro e vaca, respectivamente, obtidos sob o modelo animal implementado no primeiro passo da análises (i.e. o Modelo 1 descrito no presente capítulo).

Foram calculadas seis medidas de dispersão de RESm da progênie dos touros: desvio-padrão (SD), variância (VAR), log de VAR (ln), *median absolute deviation* (MAD), *Qn scale* (Qn) e *Sn scale* (Sn), de modo que, para cada medida de dispersão, foram estimados componentes de variância (VC) sob o modelo 2 descrito anteriormente, tanto por metodologia REML quanto por inferência bayesiana (NBVC).

Ao analisar os dados de dispersão segundo os diferentes estimadores sob o Modelo 2, também foram considerados dois tamanhos mínimos de progênie para inclusão dos touros nas análises, caracterizando dois conjuntos de dados de dispersão, referidos, posteriormente, como conjunto A e conjunto B, contendo dados de dispersão para touros com, no mínimo, 10 e 25 filhos, respectivamente.

De modo semelhante ao observado no item anterior, as estimativas de variância aditiva sobre a dispersão de RESm também foram bastante semelhantes pelas metodologias REML e NBVC (Tabela 14).

Ao se considerar o efeito de segregação mendeliana devido a touro nas observações de resíduo de sua progênie, obtiveram-se estimativas de variância aditiva de

maior magnitude, se comparada àquela observada quando se considerou exclusivamente o efeito residual no cálculo da dispersão da progênie dos touros (Tabela 9).

Menores valores de variância aditiva foram estimados no conjunto de dados B (maior tamanho mínimo de progênie), como também observado ao considerarem-se apenas dados de dispersão do efeito residual (Tabela 9). Apesar de ter sido considerado um menor número de famílias de touro no conjunto B, espera-se que as estimativas de dispersão residual neste conjunto sejam mais confiáveis, por terem sido obtidas com base em maior número de observações.

Estimativas de *evolubility* de magnitude elevada foram obtidas para todos os estimadores, após a transformação para escala de VAR (Tabela 15). Tal resultado indicaria boa oportunidade para incremento na uniformidade de GND por meio de seleção. Todavia, tais resultados devem ser interpretados com cautela, uma vez que tanto os procedimentos de transformação quanto o fato de não terem sido considerados efeitos ambientais sobre a dispersão residual de GNDa (e.g. GC) podem ter contribuído para que se superestimassem os valores de variância aditiva e de *evolubility*.

De modo geral, os valores de *evolubility* estimados neste estudo encontram-se acima do que tem sido relatado por outros autores (e.g. MULDER et al., 2007; WOLC et al., 2009), apesar destes parâmetros não serem plenamente comparáveis, uma vez que a presente definição de resíduo é diferente daquela considerada naqueles estudos.

Tabela 14 - Estimativas de variância aditiva sobre a dispersão dos resíduos (considerando efeito de segregação mendeliana devido a touro), obtidas sob diferentes estimadores e procedimentos de estimação para dois conjuntos de dados de dispersão (A e B)

Estimador	REML		NBVC	
	A	B	A	B
SD	18,26	16,78	18,40	17,00
VAR	15005	13606	14272	13268
LN	0,419	0,353	0,440	0,368
MAD	9,33	6,56	9,20	5,84
Qn	26,51	18,80	26,24	18,60
SN	21,19	14,27	22,28	14,40

A = mínimo de 10 filhos por touro; B = mínimo de 25 filhos por touro

Ao se comparar as estimativas de *evolubility* obtidas sem o efeito de segregação mendeliana devido a touro (Tabela 10) com aquelas obtidas considerando-se tal efeito

(Tabela 15), percebe-se que os valores deste parâmetro sob as duas definições de resíduo foram bastante próximos para um mesmo estimador.

Tabela 15 - Estimativas de *evolubility* para variância residual de GNDa(considerando efeito de segregação mendeliana devido a touro), obtidas considerando diferentes estimadores e procedimentos de estimação para dois conjuntos de dados de dispersão (A e B)

Estimador	REML		NBVC	
	A	B	A	B
SD	56,05%	53,27%	56,26%	53,62%
VAR	49,72%	47,32%	48,49%	46,73%
LN	64,76%	59,41%	66,33%	60,66%
MAD	45,13%	37,59%	44,82%	35,47%
Qn	65,72%	57,85%	65,38%	57,54%
SN	65,71%	53,77%	67,38%	54,02%

A = mínimo de 10 filhos por touro; B = mínimo de 25 filhos por touro

Tais resultados indicam que, apesar de, provavelmente, um efeito de escala ter levado a maiores estimativas de variância aditiva para variância residual ao se considerar a segregação mendeliana devida a touro, praticamente não houve mudança na *evolubility*, uma vez que a dispersão média calculada também apresentou maior magnitude. Sendo assim, pode-se esperar que a consideração desta definição alternativa de resíduo (RESm) não implique em incremento na resposta esperada à seleção para maior uniformidade de GND, não tendo sido encontrada, sob esta abordagem, evidência de algum efeito de prepotência dos touros sobre a variabilidade de GND.

Todavia, seria aconselhável que futuros estudos tratassem desta questão empregando outras abordagens para a predição do efeito de segregação mendeliana devido a touro, uma vez que a presente definição é influenciada em alguma medida pelos resíduos das observações e pode não ser a mais adequada para estudo da prepotência dos touros.

### 3.4. CONCLUSÕES

Há indícios de ação genética aditiva sobre a dispersão residual no ganho de peso do nascimento à desmama (GND) da progênie de touros Nelore, indicando oportunidade para incremento na uniformidade desta característica por meio de seleção, sobretudo ao considerarem touros com razoável número de filhos.

Não houve indícios de que a seleção para menor dispersão residual de GND implique em menor resposta para a média desta característica, nem de que efeitos de prepotência dos touros atuem sobre a dispersão do GND de sua progênie.

Futuros estudos devem ser conduzidos para validação da presente abordagem por meio de simulação, bem como para investigar modelos que permitam considerar simultaneamente efeitos fixos e aleatórios sobre a dispersão residual de GND.

### 3.5. REFERÊNCIAS

- BIJMA, P. Quantitative genetic models for social interaction and GxE and inherited variability. Quantitative Genetic Theory and Analysis (Course Notes), Armidale Animal Breeding Summer Course, University of New England, Armidale, 2009.
- BOLDMAN, K.G.; KRIESE, L.A.; VAN VLECK, L. D.; VAN TASSEL, C. P.; KACHMAN, S. D. 1995. A manual for use for MTDFREML- a set of programs to obtain of variance and covariances. Lincoln: Department of Agriculture /ARS. 120p.
- CALO, L.L.; MCDOWELL, R.E., VAN VLECK, L.D., MILLER, P.D. Genetic aspects of beef production among Holstein-Friesians pedigree selected for milk production. **Journal of Animal Science**, Savoy, v.37, pp. 676–682, 1973.
- GUERRA, J.L.L.; FRANKE, D.E., BLOUIN, D.C. Genetic parameters for calving rate and calf survival from linear, threshold, and logistic models in a multibreed beef cattle population. **Journal of Animal Science**, v. 84, pp. 3197-3203, 2006.
- GUTIERREZ, J. P.; NIETO, B.; PIQUERAS, P.; IBAÑEZ, N.; SALGADO, C. Genetic parameters for canalisation analysis of litter size and litter weight traits at birth in mice. **Genetics Selection Evolution**, v. 38, pp. 445–462, 2006.
- HEYDARPOUR, M.; SCHAEFFER, L.R.; YAZDI, H. Influence of population structure on estimates of direct and maternal parameters. **Journal of Animal Breeding and Genetics**, Berlim, n 125, pp. 89-99, 2008.
- HOHENBOKEN, W. D. The manipulation of variation in quantitative traits: A review of possible genetic strategies. **Journal of Animal Science**, Savoy, n. 60, pp. 101-110, 1985.
- HOULE, D. Comparing evolvability and variability of quantitative traits. **Genetics**, Bethesda, v. 130, pp. 195-204, 1992.
- HUBY M.; GOGUÉ, J.; MAIGNEL, L.; BIDANEL, J.P. Corrélations génétiques entre les caractéristiques numériques et pondérales de la portée, la variabilité du poids des porcelets et leur survie entre la naissance et le sevrage. **Journées Recherche Porcine**, v. 35, pp. 293–300, 2003.
- IBÁÑEZ-ESCRICHE, N.; VARONA, L.; SORENSEN, D.; NOGUERA, J.L. A study of heterogeneity of environmental variance for slaughter weight in pigs. **Animal**, v. 2, pp.19-26, 2008.

- LITTELL, R.C., MILLIKEN, G.A., SHROUP, W.W.; WOLFINGER, R.D. SAS System for Mixed Models. Cary: SAS Institute, 1996.
- LYNCH, M.; WALSH, B. **Genetics and Analysis of Quantitative Traits**. Sunderland: Sinauer Associates, 1998. 980 p.
- MARCONDES, C.R.; BERGMANN, J.A.G.; ELER, J.P.; FERRAZ, J.B.S.; PEREIRA, J.C.C.; PENNA, V.M. Análise de alguns critérios de seleção para características de crescimento na raça Nelore. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v. 52, pp. 83-89, 2000.
- MULDER, H. A., HILL, W. G.; VEREIJKEN, A.; VEERKAMP, R. F. Estimation of genetic variation in residual variance in female and male broiler chickens. **Animal**, v. 3, pp. 1673-1680, 2009.
- MULDER, H. A.; BIJMA, P.; HILL, W. G. Prediction of Breeding Values and Selection Responses With Genetic Heterogeneity of Environmental Variance. **Genetics**, Bethesda, v.175, pp. 1895-1910, 2007.
- MULDER, H. A.; BIJMA, P.; HILL, W. G. Selection for uniformity in livestock by exploiting genetic heterogeneity of residual variance. **Genetics Selection Evolution**, v. 40, n. 1, pp. 37-59, 2008.
- ORTIZ PEÑA, C.D.; CARVALHEIRO, R.; QUEIROZ, S.A.; FRIES, L.A. Comparison of selection criteria for pre-weaning growth traits of Nelore cattle. **Livestock Production Science**, v. 86, pp.163-167, 2004.
- ROS, M.; SORENSEN, D.; WAAGEPETERSEN, R.; DUPONT-NIVET, M.; BONNET, J. C.; MALLARD, J. Evidence for Genetic Control of Adult Weight Plasticity in the Snail *Helix aspersa*. **Genetics**, v. 168, pp. 2089-2097, 2004.
- ROUSSEEUW, P. J.; CROUX C. Alternatives to the Median Absolute Deviation. **Journal of the American Statistical Association**, Alexandria, v. 88, n. 424, p. 1273-1283, 1993.
- ROWE, S. J.; WHITE, I. M.S., AVENDAÑO, S.; HILL, W. G. Genetic heterogeneity of residual variance in broiler chickens. **Genetics Selection Evolution**, Les Ulis, n. 38, pp. 617-635, 2006.
- SOERENSEN, D.; WAAGEPETERSEN, R. Normal linear models with genetically structured residual variance heterogeneity: a case study. **Genetics Research**, Cambridge, v. 82, pp. 207–222, 2003.

STATISTICAL ANALYSIS SYSTEM - SAS. SAS Institute Inc. 2002. SAS OnlineDoc® 9. Cary, NC: SAS Institute Inc.

WOLC, A. WHITE, I. M. S.; AVENDAÑO, S.; HILL, W. G. Genetic variability in residual variation of body weight and conformation scores in broiler chickens. **Poultry Science**, n.88, pp. 1156-1161, 2009.

WOLFINGER, R. D.; KASS, R. E. Nonconjugate Bayesian analysis of variance components models. **Biometrics**, n. 56, pp. 768-774, 2000.

#### **CAPÍTULO 4. VALIDAÇÃO DA ABORDAGEM EM DOIS PASSOS PARA ESTIMAÇÃO DE PARÂMETROS GENÉTICOS PARA DISPERSÃO RESIDUAL**

**RESUMO** – Este estudo de simulação foi desenvolvido com o objetivo de validar uma abordagem em dois passos para estimação de parâmetros genéticos para a dispersão residual de características em que maior uniformidade de produção é desejada. Foram geradas 30 populações com estrutura similar àquela encontrada em dados de bovinos de corte, cada uma composta por 125.000 animais distribuídos em 4.000 grupos de contemporâneos (GC) e pertencentes a 5 gerações. Foram geradas observações fenotípicas ( $Y$ ) para cada indivíduo, com base em parâmetros genéticos simulados para a média e a variância residual de  $Y$ . No primeiro passo foram estimados, sob um modelo animal, as soluções para os diferentes efeitos simulados sobre a média, permitindo a predição dos resíduos ( $\hat{e}$ ) associados a cada observação. No segundo passo, a análise de medidas de dispersão residual na progênie de touros por meio de modelo contemplando efeito aditivo de touro (modelo DIS) foi comparada com a análise do logaritmo natural de  $(\hat{e})^2$  por meio de um modelo touro (modelo ALT). Identificou-se superioridade do modelo ALT sobre o modelo DIS.

**Palavras-chave:** evolvability, acurácia de predição, variância residual

#### 4.1. INTRODUÇÃO

Como meio de validar a abordagem em dois passos descrita no capítulo 3 deste trabalho foi realizado um estudo de simulação baseado no Método de Monte Carlo, de modo a verificar a confiabilidade das estimativas de componentes de variância no segundo passo das análises, bem como a acurácia das predições do mérito genético dos touros para variância residual (VE), considerando dois cenários: ausência de efeitos genéticos aditivos sobre a variância residual (modelo homoscedástico, HOM) e presença de efeitos aditivos sobre a variância residual (modelo heteroscedástico, HET).

Os resultados obtidos para estes critérios também foram comparados com aqueles obtidos sob uma abordagem alternativa, empregando procedimento similar ao aplicado nos estudos de BOLET et al. (2007), GARREAU et al. (2008) e MULDER et al. (2009), a fim de se obter resultados que fornecessem base de comparação para a abordagem desenvolvida no capítulo 3.

#### 4.2. MATERIAL E MÉTODOS

##### *Estrutura da população simulada*

Para a validação do processo de simulação estocástico, foram simuladas 30 populações (repetições) para cada cenário, de acordo com a seguinte estrutura:

Simulou-se uma geração base (G0) constituída de 1.000 touros e 25.000 vacas, os quais foram considerados como não aparentados, não submetidos a processo prévio de seleção e assumidos como uma amostra aleatória de uma população conceitualmente infinita.

O acasalamento aleatório das vacas com os touros da população base foi simulado de modo que a primeira geração (G1) fosse composta por 25.000 produtos (um produto por vaca), metade de cada sexo, os quais foram distribuídos em 800 grupos de contemporâneos (GC).

De modo a replicar uma estrutura próxima daquela comumente encontrada em programas de melhoramento de bovinos de corte, desenvolveu-se um programa em linguagem R (R DEVELOPMENT CORE TEAM, 2009) para simular, a cada geração, a distribuição aleatória dos produtos entre touros e grupos de contemporâneos e a definição dos acasalamentos, de modo a atender simultaneamente às seguintes restrições:

- i) Número de produtos por touro: mínimo de 10 e máximo de 100.
- ii) Número de produtos por GC: mínimo de 15 e máximo de 50.
- iii) Número de touros por GC: mínimo de 5 e máximo de 15.

Para simulação da geração seguinte (G2), o processo de seleção parental contemplou uma reposição anual de 20% dos reprodutores de cada sexo, consistindo na seleção fenotípica dos melhores 800 machos e 20.000 fêmeas da população base e dos melhores 200 machos e 5.000 fêmeas de G1. Semelhante processo de seleção e acasalamento foi repetido até que os animais da quinta geração (G5) fossem produzidos, totalizando 125.000 produtos e 4.000 grupos de contemporâneos, considerando-se todos os produtos simulados.

### *Modelos simulados*

Em cada repetição e cenário foi simulada uma característica fictícia Y, expressa em quilogramas (kg), para cada animal. Os programas para simular a estrutura dos dados e as observações fenotípicas foram desenvolvidos no *software* estatístico SAS (SAS INSTITUTE, 2002).

Exceto para os animais da população base, Y foi simulada de acordo com o modelo:

$$y_{ij} = b_i + Am_{ij} + e_{ij},$$

em que  $y_{ij}$  é a observação Y do animal j no grupo de contemporâneo (GC) i,  $b_i$  é o efeito do i-ésimo GC sobre a média de Y,  $Am_{ij}$  é o valor genético aditivo do animal j para a média de Y e  $e_{ij}$  é o resíduo simulado para cada observação.

Os valores de  $b_i$  foram tomados de uma distribuição normal com média 170 kg e variância de 400 kg<sup>2</sup>. Os valores genéticos dos animais da população base para a média de Y foram gerados a partir de uma distribuição normal ( $Am \sim N(0; 100 \text{ kg}^2)$ ) e os dos animais das gerações seguintes foram obtidos pela fórmula  $Am_{ij} = \frac{1}{2}(as_j + ad_j) + m_{ij}$ , onde  $as_j$  e  $ad_j$  são os valores genéticos verdadeiros do pai e da mãe do animal j e  $m_{ij}$  é o efeito da segregação mendeliana do animal j, tomado aleatoriamente de uma distribuição normal  $N(0; 50 \text{ kg}^2)$ .

Os resíduos  $e_{ij}$  de cada observação foram simulados seguindo o modelo básico:

$$e_{ij} = \chi * \sqrt{(400 + Av_{ij})},$$

em que  $\chi$  é um termo tomado aleatoriamente de uma distribuição normal padrão, 400 é valor simulado para média da variância residual (400 kg<sup>2</sup>) e  $Av_{ij}$  é o valor genético do animal  $j$  para variância residual.

Para cada animal, o valor genético para a variância residual foi obtido com base na fórmula:

$$Av_{ij} = \frac{1}{2}(avs_j + avd_j) + mv_{ij},$$

onde  $avs_j$  e  $avd_j$  são os valores genéticos verdadeiros do pai e da mãe do animal  $j$  para a variância residual (VE) de  $Y$ . Para os animais da geração base, os valores de  $avs_j$  e  $avd_j$  foram amostrados de uma distribuição normal  $N(0; \sigma^2_{AV})$ , sendo  $\sigma^2_{AV}$  a variância aditiva sobre VE (variável para cada cenário simulado). O termo  $mv_{ij}$  corresponde ao efeito da segregação mendeliana sobre VE para o animal  $j$ , tomado aleatoriamente de uma distribuição normal  $N(0; \frac{1}{2} \sigma^2_{AV})$ .

Os parâmetros usados na simulação para a média de  $Y$  reproduziram uma característica com herdabilidade ( $h^2$ ) igual a 0,20 e, inicialmente, seguindo uma distribuição normal com média zero e variância de 900 kg<sup>2</sup> ( $N(0;500)$ ).

### *Cenários simulados e critérios de comparação*

Três diferentes cenários para a variância residual de  $Y$  foram simulados, a saber:

HOM: Cenário homoscedástico, em que não foram simulados efeitos aditivos sobre a variância residual de  $Y$  (i.e.  $Av = 0$  para todos os animais);

HET1: Cenário heteroscedástico, em que a variância aditiva sobre a variância residual de  $Y$  foi simulada de modo a perfazer um coeficiente de *evolability* de 15% para a variância residual (i.e.  $\sigma^2_{AV} = 3600$ ).

HET2: Cenário heteroscedástico, em que a variância aditiva sobre a variância residual de  $Y$  foi simulada de modo a perfazer um coeficiente de *evolability* de 20% para a variância residual (i.e.  $\sigma^2_{AV} = 6400$ ).

Para verificar se os cenários seriam realmente comparáveis entre si e validar o processo de simulação, foram calculadas as correlações de Pearson entre os efeitos simulados sobre a média de  $Y$  e suas respectivas estimativas, bem como para os componentes de variância estimados no primeiro passo das análises.

A confiabilidade dos diferentes estimadores de variância aditiva sobre VE foi verificada por meio das médias das estimativas obtidas para  $\sigma^2_{Av}$  nos diferentes cenários, contrastando-as com os respectivos valores simulados.

Como meio de caracterizar a acurácia de predição sob cada procedimento/estimador de dispersão no segundo passo, foi calculada, para cada população, a correlação de Pearson entre os valores genéticos simulados e preditos da variância residual de Y.

Foram obtidas as médias de viés e erro quadrático médio (MSE) das diferenças esperadas na progênie para variância residual (DEPv), considerando-se os touros presentes na matriz de parentesco neste passo da avaliação. Como os valores verdadeiros de Av foram simulados na escala de variância, os valores estimados para as DEPs foram transformados para esta escala, a fim de tornar comparáveis os valores obtidos sob diferentes estimadores, conforme procedimento descrito no capítulo 3 deste trabalho.

Os valores obtidos para a correlação entre valores genéticos simulados (Av) e preditos ( $\hat{A}v$ ) foram comparados com a acurácia esperada de predição ( $r_{\hat{A}v, Av}$ ), calculada segundo a fórmula proposta por MULDER et al. (2007) para obtenção de uma estimativa aproximada deste parâmetro:

$$r_{\hat{A}v, Av} \approx \frac{rh_v}{\sqrt{t_v + (1-t_v)/N}},$$

em que  $r$  é o coeficiente de parentesco entre o touro candidato e os animais com observações de resíduo,  $t_v$  é o coeficiente de correlação intra-classe,  $h_v$  é a raiz quadrada da herdabilidade da variância residual e  $N$  é o número de observações.

#### *Estimação de componentes de variância para a média e variância residual da característica simulada*

Para cada população simulada, empregou-se a abordagem em dois passos descrita no capítulo 3 deste estudo para a estimativa dos componentes de variância associados aos efeitos sobre a média e a variância residual de Y.

No primeiro passo, estimativas REML dos componentes associados à média de Y foram obtidas, por meio de um modelo animal considerando os efeitos de GC e de animal.

As soluções obtidas no primeiro passo para tais efeitos foram utilizadas para prever os resíduos ( $\hat{\epsilon}_{ij}$ ) associados a cada observação.

Seis medidas de dispersão dos resíduos preditos da progênie de cada touro foram calculadas por meio do procedimento UNIVARIATE do pacote estatístico SAS (SAS INSTITUTE, 2002): desvio-padrão (SD), variância (VAR), log de VAR (ln), *median absolute deviation* (MAD), *Qn scale* (Qn) e *Sn scale* (Sn).

Para cada cenário e estimador de dispersão residual, obtiveram-se estimativas da variância aditiva sobre a dispersão residual na escala do respectivo estimador, usando o software MTDFREML (Boldman et al., 1995) e o seguinte modelo (referido posteriormente como modelo DIS):

$$d_{ij} = \mu + g_i + \epsilon_{ij},$$

em que  $d_{ij}$  é a estimativa de dispersão residual da progênie do touro  $j$ ,  $\mu$  é a dispersão residual média,  $g_i$  é o efeito aleatório associado ao pai  $i$  do touro  $j$  e  $\epsilon_{ij}$  é um erro aleatório associado a cada estimativa de dispersão. Assumiu-se que as estimativas de variância de pai de touro sob este modelo seriam iguais a 1/16 da variância aditiva para dispersão residual ( $\sigma^2_{AV}$ ).

#### *Modelo alternativo para o segundo passo das análises*

Um modelo alternativo (ALT) também foi investigado no segundo passo das análises, como meio de fornecer uma base de comparação para os resultados obtidos no presente estudo, para cada cenário. Optou-se pela implementação de modelo similar ao utilizado por BOLET et al. (2007) e WOLC et al. (2009), inclusive pelo fato do estimador referido como ln no capítulo 3 guardar similaridade de escala com o utilizado neste estudo e ter sido identificado como tendo as distribuições mais próximas da normalidade, sendo esperado que os dados de dispersão segundo este estimador sejam mais adequados para obtenção de estimativas REML de componentes de variância no segundo passo das análises.

Utilizaram-se os resíduos preditos  $\hat{\epsilon}_{ij}$  para calcular o valor da medida  $\ln((\hat{\epsilon}_{ij})^2)$  associada a cada observação. Tal modelo assume o logaritmo natural do quadrado do resíduo como uma medida da variabilidade residual e possui a vantagem de considerar efeitos ambientais sobre a variância residual ao nível de cada observação.

As observações de  $\ln((\hat{e})^2)$  foram analisadas sob um modelo touro considerando efeitos de GC e touro, de modo a se obterem estimativas REML para os componentes de variância, bem como predições do mérito dos touros para a variância residual de Y.

O modelo ALT pode ser resumido pela equação:

$$\ln[(\hat{e}_{ij})^2] = Xgc + Zs + \varepsilon, \text{ (Modelo ALT),}$$

em que  $\ln[(\hat{e}_{ij})^2]$  é o vetor com as observações do logaritmo natural do quadrado do resíduo de GNDa para cada filho dos touros,  $gc$  é o vetor de efeitos fixos de grupo de contemporâneos,  $s$  é o vetor de efeito aditivo de touro,  $\varepsilon$  é um vetor de efeitos residuais (no segundo passo) e  $X$  e  $Z$  são as matrizes de incidência que relacionam os elementos de  $gc$  e  $s$  a  $\ln[(\hat{e}_{ij})^2]$ , respectivamente.

No modelo ALT assumiu-se  $s$  e  $\varepsilon$  normalmente distribuídos, i.e.  $s \sim N(0, A\sigma_s^2)$  e  $\varepsilon \sim N(0, I\sigma_e^2)$ , com  $A$  sendo a matriz de parentesco entre touros, baseada na informação de pai e avô materno dos touros e  $I$  é uma matriz identidade. É importante frisar que sob este modelo alternativo, a medida de dispersão residual foi associada diretamente ao efeito de touro (e não de pai de touro), de modo que se assumiu como sendo iguais a  $1/4$  da variância aditiva as estimativas obtidas para variância de touro.

### 4.3. RESULTADOS E DISCUSSÃO

Na tabela 16, são apresentadas as estatísticas referentes às correlações de Pearson entre os efeitos simulados sobre a média de Y e as respectivas predições/estimativas obtidas após o primeiro passo das análises.

Pode-se notar razoável concordância entre os cenários quanto à correspondência entre os efeitos simulados e preditos sobre a média de Y, tendo-se obtido altas correlações médias ( $> 0,9$ ) para os efeitos de GC e resíduo. Para o valor genético para a média de Y ( $A_m$ ), a correlação média, foi menor ( $\cong 0,65$ ), porém de mesma magnitude considerando-se os diferentes cenários.

As estimativas dos componentes de variância obtidas para os diferentes cenários no primeiro passo também apresentaram razoável concordância, com valores ligeiramente superiores para a variância residual média nos cenários em que foi simulada

heterogeneidade genética sobre VE (Tabela 17). Conseqüentemente, a variância aditiva e a herdabilidade foram ligeiramente inferiores nestes cenários.

Tabela 16 - Estatísticas referentes às correlações de Pearson entre os efeitos simulados e preditos/estimados sobre a média de Y, sob três cenários para variância residual

Variável	Cenário	N	Média	SD	Min	Max
Am	HET2	30	0,6539	0,007	0,6355	0,6647
	HET1	30	0,6522	0,007	0,6353	0,6645
	HOM	30	0,6463	0,007	0,6325	0,6596
egc	HET2	30	0,9807	0,001	0,9794	0,9818
	HET1	30	0,9809	0,001	0,9796	0,9819
	HOM	30	0,9800	0,001	0,9789	0,9829
res	HET2	30	0,9065	0,001	0,9052	0,9082
	HET1	30	0,9063	0,001	0,9049	0,9083
	HOM	30	0,9057	0,001	0,9043	0,9074

Am = valor genético para a média de Y; egc= efeito de grupo de contemporâneos; res = resíduo

Apesar de estas diferenças terem sido apontadas como significativas pelo teste de Tukey, pode-se considerar que os resultados estão dentro do esperado, uma vez que, dada a magnitude da variância aditiva sobre VE ( $\sigma^2_{AV}$ ), a dispersão residual média simulada foi realmente superior sob HET1 e HET2 (Tabela 18). Além disso, o valor obtido para  $h^2$  média nos três cenários implicaria no emprego de razões de variância (ve/va) praticamente idênticas nas equações de modelos mistos do primeiro passo das análises.

Tabela 17 - Estatísticas descritivas referentes às estimativas de componentes de variância no primeiro passo das análises, sob três cenários para variância residual

Variável	Cenário	Média*		SD	Min	Max
va	HET2	100,1143	<sup>B</sup>	2,9195	89,7376	104,8942
	HET1	99,9301	<sup>B</sup>	3,6752	93,5210	109,0738
	HOM	104,5536	<sup>A</sup>	3,6181	98,2574	112,4748
ve	HET2	404,9661	<sup>A</sup>	3,3328	399,1758	412,8699
	HET1	402,9722	<sup>A</sup>	3,4039	396,2832	410,7598
	HOM	397,2606	<sup>B</sup>	3,1875	390,6858	403,1282
$h^2$	HET2	0,1982	<sup>B</sup>	0,0056	0,1785	0,2076
	HET1	0,1987	<sup>B</sup>	0,0069	0,1855	0,2158
	HOM	0,2083	<sup>A</sup>	0,0068	0,1967	0,2229

\*Valores médios referentes a 30 repetições. Valores seguidos por diferentes letras diferem entre si ( $P < 0,01$ ) pelo teste de Tukey; va=variância aditiva sobre a média de Y; ve=variância residual;  $h^2$  = herdabilidade de Y.

Tabela 18 - Média e desvio-padrão (SD) da dispersão dos resíduos simulados (sim) e preditos (pred) sob três cenários de variância residual, calculada por meio de diferentes estimadores

Estimador	Tipo de resíduo	HET2		HET1		HOM	
		Média	SD	Média	SD	Média	SD
VAR	pred	336,564	75,778	334,561	72,950	326,540	66,680
	sim	405,264	93,986	403,407	89,470	399,814	81,256
SD	pred	18,231	2,048	18,184	1,976	17,976	1,841
	sim	19,998	2,313	19,963	1,976	19,892	2,208
ln	pred	5,793	0,227	5,789	0,220	5,767	0,208
	sim	5,978	0,234	5,975	0,224	5,975	0,207
MAD	pred	12,094	2,088	12,074	2,047	12,000	1,992
	sim	13,237	2,329	13,256	2,273	13,280	2,201
Qn	pred	19,245	2,684	19,225	2,618	19,056	2,489
	sim	21,090	2,996	21,099	2,920	21,091	2,749
Sn	pred	18,114	2,592	18,085	2,531	17,951	2,422
	sim	19,855	2,905	19,856	2,805	19,864	2,675

Tais resultados indicariam que a diferente estrutura de variância residual em cada cenário não afetou a acurácia de predição no primeiro passo da avaliação, dando suporte para que as diferenças observadas entre cenários no segundo passo das análises sejam assumidas como decorrentes de diferenças entre as estruturas da variância residual de Y que caracterizam cada cenário.

Na tabela 19, apresentam-se as médias dos valores de assimetria e curtose em cada cenário, para as distribuições dos diferentes estimadores de dispersão residual utilizados no segundo passo deste estudo. Valores próximos a zero dão suporte para distribuições aproximadamente normais das observações, pressuposto necessário sob o procedimento de estimação por REML.

De modo geral, a dispersão dos resíduos simulados e a dos resíduos preditos apresentaram distribuições com assimetria e curtose bem próximas, para cada estimador e cenário. Dentre os diferentes estimadores estudados, SD e Sn foram aqueles com distribuições mais próximas da normalidade, confirmando a situação que havia sido observada para os dados de uma população real no Capítulo 3 deste estudo (exceto no caso de ln, que apresentou maior desvio da normalidade no conjunto de dados simulados).

Na tabela 20, são apresentadas as médias das estimativas de variância aditiva sobre a variância residual. Após a transformação das estimativas de cada estimador para a escala da variância de Y, percebe-se que a maioria deles tendeu a subestimar este parâmetro em presença de heterogeneidade genética sobre VE (i.e. HET2 e HET1, Tabela

20). Os estimadores Qn e Sn, por outro lado, tenderam a superestimar este parâmetro sob o cenário HET1, o que também foi observado no caso de Qn no cenário HET2.

Os resultados obtidos pelo procedimento alternativo (ALT), também indicaram a tendência de subestimar  $\sigma^2_{AV}$ , apesar destes dois tipos de abordagem considerarem diferentes observações e diferente matriz de parentesco no segundo passo das análises.

Tabela 19 - Médias de assimetria (Assim.) e curtose (Curt.) da dispersão dos resíduos simulados (sim) e preditos (pred) sob três cenários de variância residual, obtida por meio de diferentes estimadores

Estimador	Tipo de resíduo	HET2		HET1		HOM	
		Assim.	Curt.	Assim.	Curt.	Assim.	Curt.
VAR	pred	0,682	1,830	0,716	2,169	0,561	1,738
	sim	0,697	1,833	0,707	2,349	0,549	1,832
SD	pred	0,178	1,026	0,191	1,219	0,059	1,292
	sim	0,192	0,952	0,185	1,155	0,063	1,192
ln	pred	-0,342	1,362	-0,324	1,332	-0,478	1,952
	sim	-0,326	1,247	-0,336	1,362	-0,448	1,700
MAD	pred	0,259	1,182	0,221	1,176	0,189	1,249
	sim	0,316	1,189	0,251	1,294	0,211	1,213
Qn	pred	0,814	2,605	0,897	3,148	0,864	3,010
	sim	0,815	2,375	0,896	3,169	0,877	2,973
Sn	pred	0,224	1,397	0,189	1,476	0,141	1,520
	sim	0,245	1,232	0,209	1,554	0,146	1,491

Tabela 20 - Médias das estimativas de variância aditiva\* sobre a variância residual ( $\sigma^2_{AV}$ ) e valores verdadeiros simulados, para diferentes estimadores de dispersão residual e cenários de estrutura de variância residual

Modelo DIS (Estimador)	Valores estimados			Valores simulados		
	HET2	HET1	HOM	HET2	HET1	HOM
VAR	4478	2751	584	6152	3657	0
SD	4124	3253	715	6152	3657	0
ln	4322	2912	645	6152	3657	0
MAD	1975	2211	1086	6152	3657	0
Qn	8603	6706	3667	6152	3657	0
Sn	3821	4034	2010	6152	3657	0
Modelo ALT	3827	2307	477	6152	3657	0

\*Todas as estimativas foram transformadas para a escala da variância (VAR), com base nos procedimentos descritos no item anterior

Outro resultado que chama a atenção foi a obtenção de médias diferentes de zero no cenário homoscedástico (HOM) para todos os estimadores. Uma possível explicação

para tal resultado poderia ser a presença de touros com progênie de pequeno tamanho, conforme também discutido por HILL (2004), ao investigar, via simulação de Monte Carlo, a estimativa sob modelo touro de um parâmetro similar a  $\sigma^2_{Av}$ .

A proporção de repetições com estimativas não-nulas para  $\sigma^2_{Av}$  (definidas como sendo maiores do que 1% do valor verdadeiro) é apresentada para cada situação na Tabela 21. Apesar de uma maior proporção de estimativas não-nulas terem sido obtidas nos modelos heteroscedásticos (HET1 e HET2), também houve convergência com estimativas diferentes de zero para um número razoável de repetições do cenário HOM, sobretudo no caso de Qn e do modelo alternativo (ALT).

Sob esta perspectiva, os estimadores mais conservadores seriam SD e In, mas dada a magnitude do viés sob a maioria das situações observadas, seria aconselhável que este tipo de parâmetro fosse estimado apenas com dados de famílias de maior tamanho e que futuros estudos de simulação investigassem se a imposição deste tipo de restrição seria efetiva para minimizar o viés na estimativa de  $\sigma^2_{Av}$ .

Para os cenários heteroscedásticos, a despeito do viés observado nas estimativas de  $\sigma^2_{Av}$ , a correlação de Pearson entre os valores genéticos verdadeiros dos touros para variância residual (Av) e seus valores genéticos preditos ( $\hat{Av}$ ) (Tabela 23) apresentou razoável concordância com o valor esperado (Tabela 22).

Tabela 21 – Proporção de repetições com estimativas não-nulas para variância aditiva sobre a dispersão residual, para diferentes cenários e estimadores de dispersão

Modelo DIS ( <i>Estimador</i> )	Cenário		
	HET2	HET1	HOM
<i>VAR</i>	76,67%	73,33%	30%
<i>SD</i>	63,33%	63,33%	23,33%
<i>In</i>	63,33%	60%	23,33%
<i>MAD</i>	46,67%	73,33%	26,67%
<i>Qn</i>	66,67%	70%	46,67%
<i>Sn</i>	50%	80%	40%
Modelo ALT	100%	100%	56,67%

Os valores simulados para  $h^2v$  foram iguais a 0,007 e 0,012 para HET1 e HET2, respectivamente. Como estes valores são baixos, também seriam esperados baixos valores para acurácia de predição, sobretudo para touros com progênie de pequeno tamanho.

A acurácia esperada sob o modelo alternativo (ALT) seria maior do que aquela esperada com o uso de estimadores de dispersão (modelo DIS), uma vez que no modelo alternativo os resíduos são associadas diretamente ao efeito de touro (e não ao de pai do touro).

Dentre os diferentes estimadores de dispersão estudados (modelo DIS), as maiores correlações foram observadas para VAR, SD e  $ln$ , sendo os resultados para estes três estimadores praticamente idênticos em ambos os cenários. Para os estimadores robustos, os valores de  $r_{\hat{A}_v, A_v}$  foram entre 10 e 40% inferiores àqueles observados no caso de VAR, porém tal resultado também pode dever-se, pelo menos em parte, ao fato dos valores de  $A_v$  terem sido simulados originalmente na escala de VAR.

Tabela 22– Acurácias esperadas de predição em cada cenário, calculadas em função dos valores simulados para herdabilidade da variância residual ( $h^2v$ ) e do número de filhos por touro (NF)

NF	HET1**		HET2**	
25	0,10	(0,21)	0,13	(0,26)
50	0,15	(0,29)	0,18	(0,36)
100	0,20	(0,39)	0,25	(0,48)

\*\*Os valores entre parênteses referem-se à acurácia esperada sob o modelo alternativo ALT.

Tabela 23 – Média da correlação de Pearson entre o valor genético verdadeiro e o valor genético predito dos touros para variância residual ( $r_{\hat{A}_v, A_v}$ ), segundo diferentes estimadores de dispersão, para diferentes classes de tamanho de progênie (N)\*

Cenário	N	Modelo DIS						Modelo ALT
		VAR	SD	$ln$	MAD	Qn	Sn	
HET1	<50	0,19	0,19	0,19	0,11	0,14	0,14	0,17
	51-100	0,19	0,19	0,19	0,11	0,17	0,15	0,25
	> 100	0,27	0,27	0,26	0,17	0,24	0,22	0,29
	<b>Média</b>	<b>0,24</b>	<b>0,23</b>	<b>0,23</b>	<b>0,14</b>	<b>0,20</b>	<b>0,18</b>	<b>0,24</b>
HET2	<50	0,22	0,22	0,21	0,13	0,18	0,16	0,25
	51-100	0,22	0,22	0,21	0,16	0,22	0,19	0,32
	> 100	0,32	0,32	0,32	0,20	0,29	0,26	0,38
	<b>Média</b>	<b>0,27</b>	<b>0,27</b>	<b>0,27</b>	<b>0,18</b>	<b>0,24</b>	<b>0,22</b>	<b>0,32</b>

\*Sob o modelo DIS, as classes foram definidas com base no número total de observações de resíduo consideradas no segundo passo das análises

Na maioria das situações, sob o modelo alternativo (ALT), os valores de  $r_{\hat{A}_v, A_v}$  foram superiores aos do modelo DIS, o que correspondeu à expectativa inicial para as predições sob estes dois tipos de modelos. Sob o cenário com maior  $h^2v$  (HET2), a correlação  $r_{\hat{A}_v, A_v}$  foi, em média, 16% superior no caso do modelo ALT.

Conforme esperado, no cenário com maior herdabilidade da variância residual ( $h^2v$ ), as correlações entre valor genético verdadeiro e predito foram de 15 a 30% superiores em relação às aquelas observadas no cenário com menor  $h^2v$  (HET1). Mesmo neste caso, os valores absolutos da acurácia de predição podem ser considerados baixos, pensando-se em termos de resposta a seleção.

MULDER et al. (2007) verificaram, por meio de simulação, que acurácias consideráveis (acima de 70%) poderiam ser obtidas para predição do  $Av$  de touros com mais de 100 filhos. Todavia, no estudo destes autores, tais níveis de acurácia foram obtidos com a simulação de maiores valores para  $h^2v$  (maiores que o dobro dos valores simulados no presente estudo).

Sendo assim, espera-se que a acurácia de seleção, com base nos modelos investigados no presente estudo, possa atingir valores razoáveis no caso de populações com  $h^2v$  mais elevada, sobretudo considerando-se os valores que têm sido relatados na literatura (entre 2 e 5%) (MULDER et al., 2007; WOLC et al., 2009).

No presente estudo, o modelo aditivo usado para simulação dos resíduos não permitiria a simulação de maiores valores de  $\sigma^2_{Av}$  (e conseqüentemente de  $h^2v$ ), mantidos os demais parâmetros. Tal fato se deve à possibilidade de valores extremamente negativos de  $Av$  levarem a valores negativos para a variância residual (o que seria intratável do ponto de vista matemático), como discutido por BIJMA (2009). Haveria a possibilidade de simular maiores valores de  $\sigma^2_{Av}$  impondo-se restrição sobre o valor mínimo de  $Av$ , todavia entendeu-se como mais prudente evitar-se esta prática, dada a possibilidade de gerar desvios da distribuição simulada em relação àquela considerada no modelo.

Na tabela 24, apresentam-se as médias obtidas para viés e erro quadrático médio dos valores genéticos para variância residual, sob os diferentes modelos de avaliação e estimadores de dispersão investigados.

Para todos estimadores de dispersão estudados, os valores de viés de predição foram bastante próximos entre si, sendo ligeiramente menores no caso de  $ln$ , nos dois cenários heteroscedásticos simulados. Ainda assim, valores consideravelmente menores de viés foram obtidos sob o modelo ALT, indicando que a estimação de  $Av$  sob este modelo seria menos tendenciosa.

Sob o modelo DIS, o menor valor médio para MSE foi observado no caso de VAR.

Menores valores de MSE também foram observados sob o modelo ALT em comparação aos modelos em que se consideraram medidas de dispersão residual, evidenciando também uma menor variância do estimador sob este modelo. Sendo assim, tanto com base no critério de viés quanto no de MSE, houve evidências de que o modelo ALT fosse mais adequado para predição do mérito genético para variância residual.

Como não foi simulado efetivamente nenhum efeito ambiental sobre a variância residual de Y, verificou-se a hipótese de que o modelo ALT sem a inclusão de efeito de GC se ajustasse melhor aos dados, porém a remoção deste efeito pouco influenciou as estimativas de variância de touro, bem como os valores de viés e MSE (dados não apresentados).

Tabela 24 – Médias de viés ( $\text{kg}^2$ ) e erro quadrático médio (MSE, em  $\text{kg}^4$ ) da diferença esperada na progênie para variância residual, sob diferentes cenários e estimadores de dispersão residual

Cenário	Estatística	Modelo DIS						Modelo ALT
		VAR	SD	In	MAD	Qn	Sn	
HET1	viés	-3,985	-4,032	-3,886	-3,998	-3,962	-4,014	-0,352
	MSE	908,24	931,57	929,12	931,35	930,84	931,33	851,99
HET2	viés	-7,033	-7,014	-6,959	-7,033	-6,898	-7,044	-0,436
	MSE	1534,51	1586,12	1563,21	1586,53	1583,36	1586,51	1417,61

Para verificar o efeito do processo de seleção fenotípica sobre as populações simuladas são apresentadas as tendências fenotípicas (Figura 3) e as tendências dos valores genéticos verdadeiros para a média e a variância residual de Y (Figura 4), nos diferentes cenários simulados.

Pode-se constatar que praticamente não houve diferenças entre os cenários em termos de média fenotípica ao longo das gerações, havendo um acréscimo em torno de 6 kg na média de Y, após 5 gerações de seleção.

Nos cenários heteroscedásticos (HET1 e HET2), a tendência dos valores genéticos verdadeiros para a média de Y ( $A_m$ ) também foi bastante próxima, o que não ocorreu no caso do valor genético médio para a variância residual, observando-se maior incremento no valor médio de  $A_v$  no cenário em que foi simulada maior variância para este efeito (HET2).

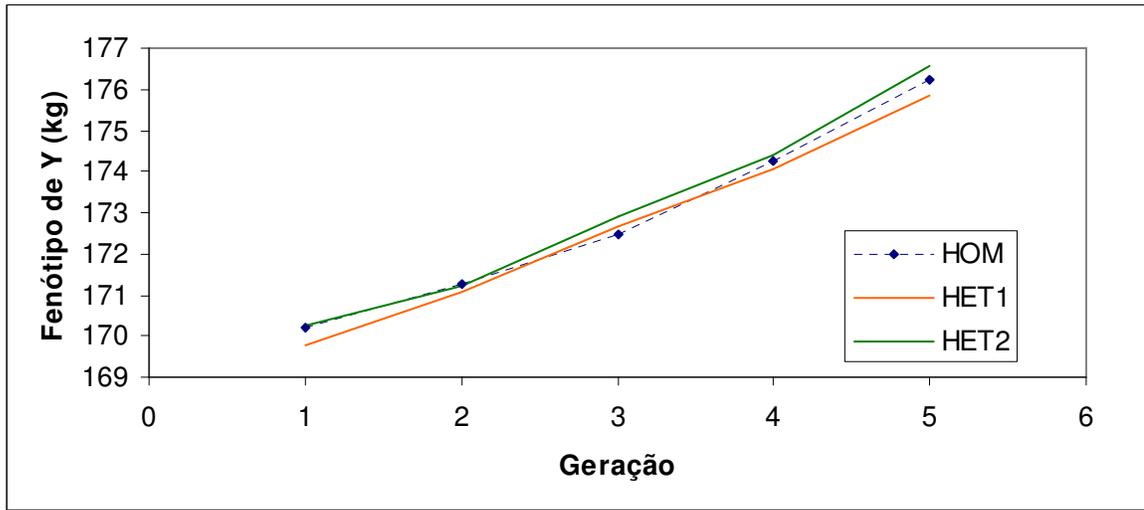


Figura 3. Tendências\* fenotípicas de Y para os diferentes cenários simulados  
\* Valores referentes à média de 30 repetições, para cada geração e cenário.

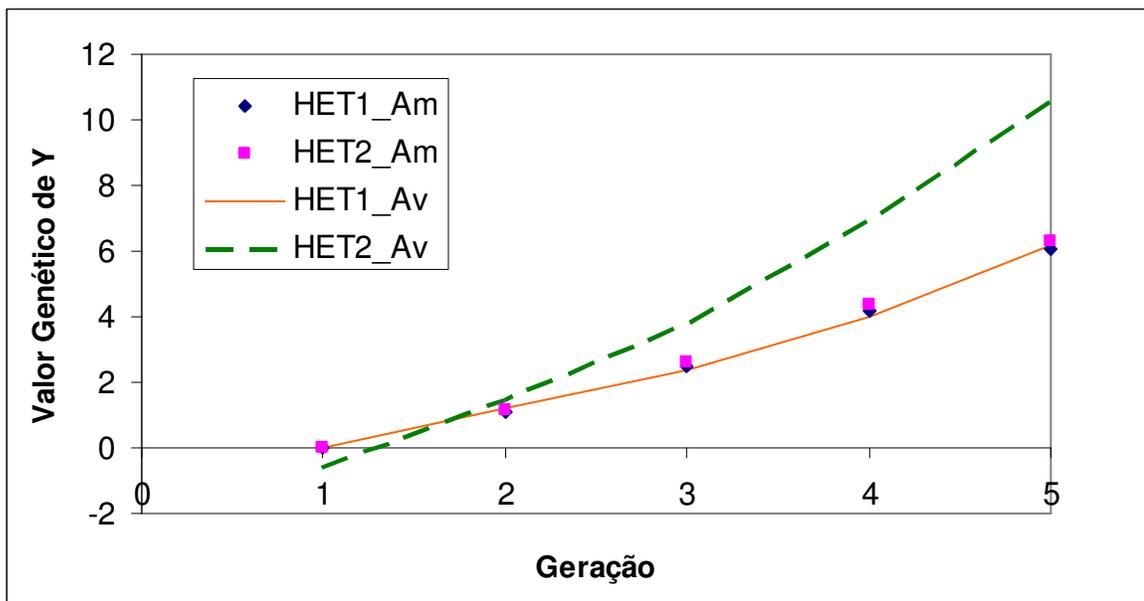


Figura 4. Tendências\* dos valores genéticos verdadeiros para a média (Am, em kg) e a variância residual (Av, em kg<sup>2</sup>) de Y, para os dois cenários com heterogeneidade genética sobre a variância residual (HET1 e HET2)

\* Valores referentes à média de 30 repetições, para cada geração e cenário.

Tais resultados confirmariam os argumentos apresentados por HILL (2004), que indicou que sob seleção intensa por truncamento os genótipos mais variáveis teriam maior probabilidade de serem selecionados. No caso do presente estudo, em que foi simulada correlação genética nula entre a média e a variância residual de Y, tal efeito de seleção indireta para maior variabilidade foi observado.

Mais estudos sobre esta questão seriam importantes para que se pudesse prever, de modo mais preciso, a resposta à seleção para maior uniformidade quando se considera simultaneamente outros critérios de seleção. Neste caso, o efeito descrito por HILL (2004), poderia contrabalancear o efeito de uma correlação genética favorável à seleção simultânea para incremento na média e redução na variabilidade (como aquela estimada no estudo do capítulo 3). Além disso, a influência de outros efeitos, como o Efeito Bulmer e uma eventual presença de genes de efeito maior, poderia tornar ainda mais complexa este tipo de predição.

As abordagens em dois passos investigadas neste estudo de simulação forneceram predições razoáveis do mérito genético dos animais para variância residual, considerando-se a magnitude simulada para a herdabilidade da variância residual, tendo sido verificada superioridade para o modelo ALT (maior acurácia e menor viés de predição) em detrimento do uso de medidas de dispersão (DIS).

Além disso, a estimação de parâmetros genéticos para variância residual em dois passos, empregada neste estudo, teria como vantagem a menor complexidade para implementação dos modelos e a possibilidade de utilização dos *softwares* usualmente empregados em melhoramento genético para estimação de parâmetros e posterior avaliação genética dos animais para uniformidade de produção.

É importante ressaltar que a validade dos resultados obtidos por meio deste estudo de simulação restringe-se a situações em que a magnitude de efeitos ambientais ou genéticos não aditivos sobre a variância residual seja negligenciável, uma vez que foram simulados apenas efeitos aditivos sobre VE. Apesar de efeitos não aditivos não poderem ser discriminados sob o modelo DIS, seria interessante investigar, em estudos futuros, as propriedades deste tipo de modelo em presença de heterocedasticidade devida a efeitos ambientais (heterogeneidade entre GCs, por exemplo), também o comparando com o modelo ALT.

Espera-se que sob maiores valores de  $h^2v$  a abordagem em dois passos investigada no presente estudo permita seleção com razoável acurácia para incremento de uniformidade, no caso de touros com progênie de grande tamanho (digamos 100 filhos). Todavia, as estimativas não-nulas obtidas para  $\sigma^2_{Av}$  sob o cenário homoscedástico simulado apontam para a necessidade de estudos futuros que permitam identificar procedimentos mais robustos para estimação deste parâmetro, bem como investigar modelos para estimação de  $\sigma^2_{Av}$  considerando, simultaneamente, efeitos ambientais e aditivos sobre a variância residual das características de produção.

#### *4.4. CONCLUSÕES*

Sob ausência de efeitos ambientais sobre a variância residual, os modelos DIS e ALT permitiriam predições do mérito genético de touros para dispersão residual com razoável acurácia, sendo este último modelo superior ao primeiro em termos de habilidade preditiva.

Futuros estudos de simulação poderiam ser conduzidos visando investigar a habilidade preditiva destes modelos sob a ação de efeitos ambientais e aditivos sobre a dispersão residual, bem como para investigar modelos mais robustos para a estimação de parâmetros genéticos para a variância residual de características de produção.

#### 4.5. REFERÊNCIAS

- BIJMA, P. Quantitative genetic models for social interaction and GxE and inherited variability. Quantitative Genetic Theory and Analysis (Course Notes), Armidale Animal Breeding Summer Course, University of New England, Armidale, 2009.
- BOLDMAN, K.G.; KRIESE, L.A.; VAN VLECK, L. D.; VAN TASSEL, C. P.; KACHMAN, S. D. 1995. **A manual for use for MTDFREML**: a set of programs to obtain of variance and covariances. Lincoln: Department of Agriculture /ARS. 120p.
- BOLET, G.; GARREAU, H.; JOLY, T.; THEAU-CLEMENT, M.; FALIERES, J.; HURTAUD, J.; BODIN, L. Genetic homogenization of birth weight in rabbits: indirect selection response for uterine horn characteristics. **Livestock Science**, v. 111, pp. 28–32, 2007.
- GARREAU, H.; BOLET, G.; LARZUL, C.; ROBERT-GRANIÉ, C.; SALEIL, G.; SANCRISTÓBAL-GAUDY, M.; BODIN, L. Results of four generations of a canalizing selection for rabbit birth weight. **Livestock Science**, v. 119, pp. 55-62, 2008.
- HILL, W. G. Heterogeneity of genetic and environmental variance of quantitative traits. **Journal of the Indian Society of Agricultural Statistics**, v. 57, pp. 49-63, 2004.
- MULDER, H. A., HILL, W. G.; VEREIJKEN, A.; VEERKAMP, R. F. Estimation of genetic variation in residual variance in female and male broiler chickens. **Animal**, v. 3, pp. 1673-1680, 2009.
- MULDER, H. A.; BIJMA, P.; HILL, W. G. Prediction of breeding values and selection responses with genetic heterogeneity of environmental variance. **Genetics**, Bethesda, v.175, pp. 1895-1910, 2007.
- R DEVELOPMENT CORE TEAM (2009). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: <http://www.R-project.org>.
- STATISTICAL ANALYSIS SYSTEM - SAS. SAS Institute Inc. 2002. SAS OnlineDoc® 9. Cary, NC: SAS Institute Inc.
- WOLC, A. WHITE, I. M. S.; AVENDANO, S.; HILL, W. G. Genetic variability in residual variation of body weight and conformation scores in broiler chickens. **Poultry Science**, n.88, pp. 1156-1161, 2009.

## **CAPÍTULO 5. COMPARAÇÃO DE MODELOS ALTERNATIVOS CONSIDERANDO HETEROGENEIDADE GENÉTICA E AMBIENTAL SOBRE A VARIÂNCIA RESIDUAL DO GANHO DE PESO DO NASCIMENTO À DESMAMA DE BOVINOS DA RAÇA NELORE**

**RESUMO** – Este estudo foi desenvolvido a fim de investigar a influência de efeitos ambientais e genéticos sobre a variância residual do ganho de peso do nascimento à desmama (GND) de bovinos Nelore, por meio de uma abordagem em dois passos. No primeiro passo foram obtidas soluções para diferentes efeitos considerados sobre a média de GND, permitindo a predição dos resíduos ( $\hat{\epsilon}$ ) associados a cada observação. No segundo passo, estimou-se a variância aditiva para a variância residual de GND por meio de diferentes modelos, assumindo-se logaritmo natural de  $(\hat{\epsilon})^2$  como um estimador da variância residual. Houve evidências de efeitos aditivos e ambientais influenciando simultaneamente a variância residual de GND, bem como de que a seleção para menor variância residual implicaria em alguma redução na resposta sobre a média desta característica. Ainda assim, considerável resposta à seleção para uniformidade de GND poderia ser esperada no caso de seleção para menor variância residual, sobretudo no caso de seleção intensa e consideração de famílias de touro de grande tamanho.

**Palavras-chave:** evolvability, heterocedasticidade, resposta correlacionada, variância residual

### 5.1.INTRODUÇÃO

Tendo em vista algumas limitações identificadas sob a abordagem em dois passos desenvolvida nos estudos dos capítulos 3 e 4 (e.g. impossibilidade de contemplar efeitos ambientais sobre a variância residual de GND e a dificuldade de ponderar diretamente o tamanho da progênie dos touros no segundo passo das análises), procurou-se desenvolver modelos alternativos que permitissem contemplar tais questões ao estimar parâmetros genéticos sobre a variância residual.

Os resultados obtidos no estudo de simulação do capítulo 4 deram suporte para que se considerasse o modelo referido como ALT como uma alternativa promissora para estudos considerando heterogeneidade genética da variância residual em bovinos de corte, sendo que modelos bastante similares a este têm sido usados em estudos desta natureza (e.g. BOLET, 2007; GARREAU et al., 2008; WOLC et al., 2009).

Este tipo de modelo permite utilizar diretamente, para cada observação, os resíduos preditos para uma determinada característica, ao assumir-se o logaritmo natural do quadrado dos resíduos como um estimador da variância residual ( $\sigma^2_E$ ). Além disso, a implementação deste modelo por uma abordagem em dois passos reduz consideravelmente a complexidade dos algoritmos de estimação comparada à de modelos implementados num único passo para o mesmo fim.

Deste modo, este estudo foi desenvolvido a fim de investigar simultaneamente a influência de efeitos ambientais (efeito de grupo de contemporâneos) e aditivos sobre a variância residual de GND.

### 5.2.MATERIAL E MÉTODOS

Utilizaram-se os mesmos dados da base Aliança Nelore, para ganho de peso do nascimento à desmama empregados nos estudos do capítulo 3 deste trabalho, compreendendo informações de desempenho de 325.432 animais, filhos de 3.339 touros e 200.601 vacas, que posteriormente foram pré-ajustadas para os efeitos fixos de idade do animal (padrão de 205 dias), idade da vaca ao parto (padrão de 7 anos) e data juliana média dentro da estação de nascimento, sendo as observações resultantes referidas como ganho do nascimento à desmama pré-ajustado (GNDa).

O primeiro passo das análises foi realizado considerando o Modelo 1 descrito em maiores detalhes no capítulo 3. Estimativas para componentes de variância e soluções para os efeitos fixos e aleatórios sobre GNDa foram obtidos por meio do software MTDFREML (BOLDMAN et al., 1995), em que se implementou o seguinte modelo animal:

$$GNDa = Xgc + Z1ad + Z2am + Z3ap + e, \text{ (Modelo 1)}$$

considerando o efeito fixo de grupo de contemporâneos (gc, definido pela concatenação das informações de rebanho, ano, estação, sexo, data de pesagem e grupo de manejo) e os efeitos aleatórios aditivo direto (ad), aditivo materno (am), de ambiente permanente devido à mãe (ap) e residual (e), sendo X, Z1, Z2 e Z3 as respectivas matrizes de incidência de cada efeito sobre GNDa.

As soluções obtidas no primeiro passo para os efeitos do Modelo 1 foram utilizadas para prever os resíduos ( $\hat{e}$ ) associados a cada observação. Em seguida, utilizaram-se os resíduos  $\hat{e}_{ij}$  para calcular o valor da medida  $\ln((\hat{e})^2)$  para cada observação de GNDa.

As observações de  $\ln((\hat{e})^2)$  foram analisadas sob diferentes modelos considerando efeitos de GC e/ou aditivos sobre a variância residual, de modo a se obterem estimativas REML (usando o *software* MTDFREML) para os componentes de variância, bem como predições do mérito dos touros para a variância residual de GNDa.

A opção pelo estudo de diferentes modelos justifica-se pelo fato dos resultados obtidos até então não terem possibilitado discriminar, simultaneamente, a magnitude de efeitos fixos e aditivos sobre a variância residual de GND. Sendo assim, na tabela 25 são apresentados os efeitos que caracterizaram cada modelo empregado para estimativa de componentes de variância para os dados de  $\ln((\hat{e})^2)$ .

Tabela 25 – Efeitos fixos (GC) e aleatórios (Touro e Animal) considerados nos diferentes modelos empregados para estimação de componentes de variância para dispersão residual de GND

Modelo	GC	Touro	Animal
LNE1	X		
LNE2		X	
LNE3			X
LNE4	X	X	
LNE5	X		X

Posteriormente, os cinco modelos estudados foram classificados por meio da estatística  $-2 \cdot \ln(L)$ , em que  $L$  é o valor maximizado da função de verossimilhança de cada modelo após atingir-se convergência (i.e. variância menor que  $10^{-9}$  para a função simplex). Também foi utilizado o critério de informação de Akaike (AIC), visando permitir a comparação entre os modelos não aninhados, penalizando os modelos com maior número de parâmetros, ( $AIC = -2 \ln(L) + 2 \cdot p$ , em que  $p$  é o número de parâmetros do modelo). Menores valores para estas duas estatísticas indicam melhor ajuste aos dados.

Os modelos aninhados também foram comparados pelo teste de razão de verossimilhança (LRT), de modo a verificar se a consideração de efeitos aditivos (LNE4 vs. LNE1, LNE5 vs. LNE1) ou de GC (LNE4 vs. LNE2, LNE5 vs. LNE3) na estimação de componentes de variância sobre a dispersão residual realmente implicou em mudanças significativas na função de verossimilhança.

As estimativas de variância aditiva sobre a variância residual ( $\sigma^2_{Av}$ ) na escala exponencial foram transformadas para a escala da variância, empregando-se os procedimentos propostos no estudo de MULDER et al. (2007). Os coeficientes de herdabilidade da variância residual ( $h^2_v$ ) e *evolubility* ( $Ev$ ) também foram calculados com base nas fórmulas propostas por estes autores.

A fórmula proposta por MULDER et al. (2007) para obtenção de estimativas da acurácia de predição do mérito dos touros para variância residual foi usada para calcular a acurácia aproximada de predição, considerando-se touros com 100 ou 200 filhos:

$$r_{\hat{A}_v, A_v} \approx \frac{1/2 * h_v}{\sqrt{t_v + (1-t_v)/N}},$$

em que  $1/2$  é o coeficiente de parentesco entre o touro candidato e os animais com observações de resíduo para GND,  $t_v$  é o coeficiente de correlação intra-classe,  $h_v$  é a raiz quadrada da herdabilidade da variância residual ( $h^2_v$ ) e  $N$  é o número de filhos por touro.

Como meio de investigar indiretamente uma eventual associação entre os valores genéticos preditos (EBVs) para a média de GND (no primeiro passo das análises) e os EBVs para a variância residual (no segundo passo das análises), calculou-se a correlação de Spearman entre estes dois tipos de EBV dos touros. A fim de considerar apenas touros com EBVs preditos com maior acurácia, foram incluídos neste cálculo apenas touros com, no mínimo, 50 filhos.

Eventuais implicações da seleção contemplando simultaneamente incremento na média e redução na variância residual de GND foram investigadas empregando-se procedimento similar àquele utilizado no estudo de NEVES et al. (2008), assumindo-se a aplicação de um índice combinando mérito para desempenho médio e uniformidade de progênie, visando retratar uma situação de uso de DEPs dos touros para variabilidade da progênie frente a diferentes condições de mercado (variação da importância relativa da uniformidade de GND).

As soluções obtidas no primeiro passo das análises para efeito aditivo direto sobre a média de GND foram usadas para calcular as DEPs dos touros para tendência central de GND (posteriormente referidas como  $DEP_A$ ), assim como as soluções obtidas sob o modelo LNE5 para efeito aditivo foram usadas para calcular DEPs para uniformidade de GND ( $DEP_U$ ), após transformação das soluções para a escala de variância conforme proposto por MULDER et al. (2007), seguida de multiplicação por -1, para expressão das  $DEP_U$  no sentido de uniformidade.

Para cada touro, calculou-se um índice (I) combinando a DEP para tendência central ( $DEP_A$ ) e a  $DEP_U$  para uniformidade, variando-se de 0 a 1 a importância relativa da uniformidade da progênie em relação ao valor de I (perfazendo 21 cenários), de modo que:

$$I = sDEP_A * (1-u) + sDEP_U * u, \text{ em que:}$$

$sDEP_A$  é a DEP padronizada para a tendência central de GNDa,  $sDEP_U$  é a DEP padronizada para uniformidade e u é a importância relativa da uniformidade.

Foram gerados três conjuntos de dados, contendo a avaliação genética dos touros com, no mínimo, 25, 50 e 100 filhos. Posteriormente, em cada conjunto de dados, os touros foram classificados em ordem decrescente e agrupados em decis, obtendo-se a média das DEPs para tendência central e uniformidade de GND em cada decil, de modo a possibilitar a comparação entre o efeito da importância relativa da uniformidade na seleção com base no índice.

Por fim, foram verificadas as tendências das médias das predições dos valores genéticos para a tendência central e para variância residual de GND, segundo o ano de nascimento dos animais, considerando as soluções obtidas com base no modelo que melhor se ajustou aos dados no segundo passo.

### 5.3. RESULTADOS E DISCUSSÃO

Inicialmente, os coeficientes de assimetria e curtose evidenciaram severos desvios da normalidade para os dados de  $\ln((\hat{\epsilon})^2)$ , mesmo sendo tais dados provenientes de um procedimento de transformação logarítmica.

MULDER et al. (2009) também relataram ocorrência similar em dados de frangos de corte, porém desenvolveram simulações que confirmaram que a transformação logarítmica seria adequada para estimar a variância genética aditiva para variância residual.

Ainda assim, considerando a magnitude do desvio da normalidade no presente conjunto de dados e sua eventual influência para a estimação de componentes de variância, optou-se pela remoção de observações discrepantes (fora do intervalo compreendido por  $\text{média} \pm 3 \cdot \text{SD}$ ). Após este procedimento, os coeficientes de assimetria e curtose para  $\ln((\hat{\epsilon})^2)$  foram iguais a -0,782 e 0,515, respectivamente, obtendo-se assim uma distribuição mais próxima da normalidade.

Após a remoção de observações consideradas discrepantes, os dados de  $\ln((\hat{\epsilon})^2)$  efetivamente usados no segundo passo das análises apresentaram média igual a 3,916 e desvio-padrão (SD) igual a 2,055, obtendo-se estes valores ao considerarem-se 321.139 observações de GNDa.

Na tabela 26, são apresentados os valores das estatísticas referentes à função de verossimilhança, juntamente com os resultados para o teste LRT, com os modelos classificados segundo ordem crescente para os valores de AIC.

Com base nos valores da função de verossimilhança e de AIC, verificou-se que o modelo animal incluindo o efeito fixo de GC (LNE5) foi o que melhor se ajustou aos dados de variância residual de GND, seguido do modelo touro também considerando efeito de GC (LNE4).

A consideração de efeitos aditivos sobre a variância residual implicou em melhor ajuste aos dados e aumento significativo na função de verossimilhança, verificado com base no teste LRT (LNE5 vs. LNE1 e LNE4 vs. LNE1). Sendo assim, parece haver suporte para a ação simultânea de efeitos aditivos e ambientais sobre a variância residual da característica estudada.

Dada a considerável alteração na função de verossimilhança para modelos considerando efeito de GC sobre a variância residual comparando com modelos

considerando apenas efeitos aditivos ( LNE5 vs. LNE3 e LNE4 vs. LNE2), pode-se entender que a influência de efeitos ambientais sobre VE é de magnitude elevada no caso dos dados de GNDa (Tabela 26). Tal resultado poderia ser uma justificativa para os valores elevados de *evolvability* que foram estimados quando se ignorou este tipo de efeito (i.e. resultados do Capítulo 3 deste estudo), fornecendo indícios de que a variância aditiva sobre a dispersão residual possa ter sido superestimada naquele caso.

Na tabela 27, apresentam-se as estimativas de variância aditiva na escala exponencial e após sua transformação para a escala de variância, assim como os valores obtidos para  $h^2v$ , *evolvability* e a acurácia esperada de predição do mérito dos touros.

Tabela 26 – Resultados referentes à função de verossimilhança ( $-2\ln(L)$ ), critério de informação de Akaike (AIC) e teste de razão de verossimilhança (LRT) para os diferentes modelos e classificação (Rank) do modelo segundo melhor ajuste aos dados

Modelo	$-2\ln(L)$	AIC	LRT	Rank
LNE5	765.462	765.466	40 <sup>*,a</sup> (14.675 <sup>*,b</sup> )	1
LNE4	765.486	765.490	17 <sup>*,c</sup> (15.189 <sup>*,d</sup> )	2
LNE1	765.502	765.504	-	3
LNE3	780.137	780.139	-	4
LNE2	780.675	780.677	-	5

LNE1- modelo considerando apenas efeito de GC sobre a variância residual de GND

LNE2- modelo considerando apenas efeito de touro sobre a variância residual de GND

LNE3- modelo considerando apenas efeito de animal sobre a variância residual de GND

LNE4- modelo considerando efeito de touro e GC sobre a variância residual de GND

LNE5- modelo considerando efeito de animal e GC sobre a variância residual de GND

<sup>\*</sup>P<0,001. Comparações: <sup>a</sup>LNE5 vs. LNE1, <sup>b</sup>LNE5 vs. LNE3, <sup>c</sup>LNE4vs. LNE1, <sup>d</sup>LNE4 vs. LNE2

Tabela 27 – Estimativas de parâmetros genéticos sobre a variância residual de GND

Modelo	$\sigma^2_{Av,exp}$	$\sigma^2_{Av,add}$	Ev	$h^2v$	Acc1	Acc2
LNE5	0,02318	2.956	26,84%	0,0055	0,349	0,466
LNE4	0,00824	1.027	15,82%	0,0019	0,216	0,298
LNE3	0,39074	91.489	149,30%	0,1146	0,864	0,925
LNE2	0,39988	95.173	152,28%	0,1176	0,867	0,926

$\sigma^2_{Av,exp}$  e  $\sigma^2_{Av,add}$  = estimativas de variância aditiva sobre a variância residual, na escala exponencial e transformadas para escala de variância, respectivamente. Ev=evolvability, em %.  $h^2v$ =herdabilidade da variância residual de GND. Acc1 e Acc2= acurácias esperadas de predição considerando touros com 100 e 200 filhos, respectivamente.

LNE2- modelo considerando apenas efeito de touro sobre a variância residual de GND

LNE3- modelo considerando apenas efeito de animal sobre a variância residual de GND

LNE4- modelo considerando efeito de touro e GC sobre a variância residual de GND

LNE5- modelo considerando efeito de animal e GC sobre a variância residual de GND

Comparando-se os modelos considerando apenas efeitos aditivos sobre a variância residual (LNE3 e LNE2) com os modelos LNE5 e LNE4, percebe-se que a não consideração de efeitos ambientais sobre VE implicou em estimativas superestimadas da variância, também implicando na superestimação de  $Ev$  e  $h^2v$ , o que fica claramente evidenciado no caso de valores de *evolvability* superiores a 100%.

Comparando-se os modelos LNE5 e LNE4, identificados como superiores em termos de ajuste aos dados, verifica-se que a estimação de  $\sigma^2_{Av}$  segundo modelo animal forneceu estimativas mais elevadas para  $h^2v$  e  $Ev$ , implicando em que a resposta esperada da seleção para incremento na uniformidade seja maior sob este modelo.

Com base nos valores de *evolvability* encontrados pode-se esperar uma resposta à seleção inferior àquela predita no Capítulo 3, todavia com magnitude ainda considerável sob algumas situações.

Outro ponto importante é que a remoção de observações discrepantes no segundo passo das análises pode ter levado à uma subestimação da variabilidade genética para variância residual, como reportado por ROWE et al. (2006). Estes autores relataram que a remoção de observações discrepantes implicou em redução nos coeficientes estimados para *evolvability* da variância residual do peso corporal de frangos de corte, sendo que tais dados também apresentaram assimetria negativa e curtose positiva, como no presente estudo.

Apesar de terem empregado este procedimento visando eliminar observações de animais que apresentaram pesos extremos em virtude de doenças ou de fatores não identificáveis, ROWE et al. (2006) indicaram que a remoção de observações discrepantes poderia ter levado a uma subestimação deste parâmetro, uma vez que famílias cujos indivíduos são mais variáveis têm maior probabilidade de terem indivíduos excluídos. Tal fato também ficou evidenciado no presente estudo, uma vez que das 3339 famílias de touro presentes no arquivo original, dados de apenas sete famílias foram considerados discrepantes.

Sendo assim, seria interessante que estudos futuros também investigassem melhor esta questão, talvez procurando identificar um ponto ótimo de *trimming* para estudos ligados a variabilidade genética sobre a variância residual, ou ainda investigassem o uso de

procedimentos de estimação em que a normalidade dos dados não fosse requerida (e.g. uso de modelos generalizados).

Conforme discutido no estudo de MULDER et al. (2008), os resultados obtidos para a acurácia esperada de predição indicam que ao se considerarem simultaneamente efeitos ambientais e aditivos sobre a variância residual há necessidade de famílias de grande tamanho, para que se possa predizer de modo acurado o mérito dos touros para dispersão, ao se discriminar adequadamente os efeitos aditivos dos efeitos ambientais.

Apesar da acurácia de predição sob modelo animal poder ter sido subestimada usando-se a fórmula proposta por MULDER et al. (2007), obteve-se estimativas de que, para touros com cerca de 200 filhos, seria esperada acurácia próxima de 50%.

Ainda assim, considerando-se um cenário de seleção intensa para menor variância residual (digamos dos 20% melhores touros) e o valor de *evolvability* encontrado sob o modelo LNE5, pode-se predizer uma redução de 6,50% e 8,75% na variância residual de GNDa após uma geração de seleção, para touros com 100 e 200 filhos, respectivamente.

De modo geral, tanto os valores de  $h^2v$  quanto de *evolvability* foram ligeiramente inferiores ao que tem sido relatado na literatura para a variância residual de características de peso corporal em frangos de corte ( ROWE et al. 2006; MULDER et al., 2009; WOLC et al., 2009). Todavia, não há estudos que permitam uma comparação mais próxima com dados de bovinos de corte. Apesar deste fato, pode-se esperar que a seleção para menor variância residual da característica investigada no presente estudo implique em resposta considerável em termos de incremento de uniformidade, principalmente, em situações de seleção mais intensa e famílias de grande tamanho.

Obtiveram-se estimativas significativas e positivas para a correlação de Spearman entre EBVs dos touros para efeito direto sobre a média de GNDa e para a variância residual sob os modelos em que se considerou efeito de GC no segundo passo, com valores iguais a 0,159 e 0,133 para LNE5 e LNE4, respectivamente. Todavia, para os modelos em que não se considerou efeito de GC no segundo passo, estimativas negativas para este tipo de correlação foram obtidas, com valores iguais a -0,173 e -0,177 para os modelos LNE3 e LNE2 respectivamente. Sob todos os modelos estudados não foi verificada nenhuma correlação significativa com o EBV para efeito materno. Tal resultado indicaria que as estimativas de correlação negativa entre média e dispersão de GNDa obtidas no estudo do

Capítulo 3 possam ter sido viesadas pela não consideração de efeitos ambientais sobre a variância residual de GNDa nos modelos considerados naquele caso.

As estimativas de correlação positiva obtidas nos modelos que melhor se ajustaram aos dados caracterizariam uma associação desfavorável entre estes dois critérios de seleção, implicando que a obtenção de alguma redução na variância residual de GNDa por meio de seleção leve a uma menor resposta sobre a sua média.

Apesar da baixa magnitude, os valores obtidos para correlação entre EBVs para média e variância residual de GNDa destoam daqueles relatados nos estudos de MULDER et al. (2009) e de WOLC et al. (2009), ambos obtidos para dados de peso corporal de frangos de corte, sendo que no primeiro estudo esta estimativa de correlação foi obtida sob um modelo bivariado (incluindo a média e o  $\ln(\hat{\sigma}^2)$ , em um processo iterativo) e, no segundo, foram obtidas estimativas de correlação entre DEPs, sendo esta abordagem mais próxima da empregada com os dados de GNDa.

IBAÑEZ-ESCRICHE et al. (2008), investigando heterogeneidade genética sobre a variância residual do peso ao abate de suínos, por meio da implementação de um modelo num único passo sob abordagem *bayesiana*, indicaram ausência de correlação (ou pequena correlação negativa) entre a média e a variância residual desta característica (i.e. o valor 0 foi compreendido pela região de credibilidade de 95% para este parâmetro). Por outro lado, alguns autores relataram correlações elevadas e positivas entre média e variância residual (ROS et al., 2004; GUTIERREZ et al. 2006).

Deste modo, dada a divergência entre os resultados obtidos em diferentes estudos para este parâmetro, é necessário investigar de modo mais aprofundado a magnitude deste tipo de correlação para características de produção em bovinos de corte, também procurando estudar a influência de efeitos de escala sobre este parâmetro e sobre as estimativas obtidas para a variância aditiva sobre VE.

Numa situação de seleção simultânea para a média e uniformidade de GND seria esperado que, considerando a seleção dos 10% melhores touros para o índice I, a obtenção de alguma resposta em termos de uniformidade (i.e. média das DEPu maior que 0) seria verificada em situações em que se atribuísse importância relativa de, no mínimo, 30% para este critério (Figura 5).

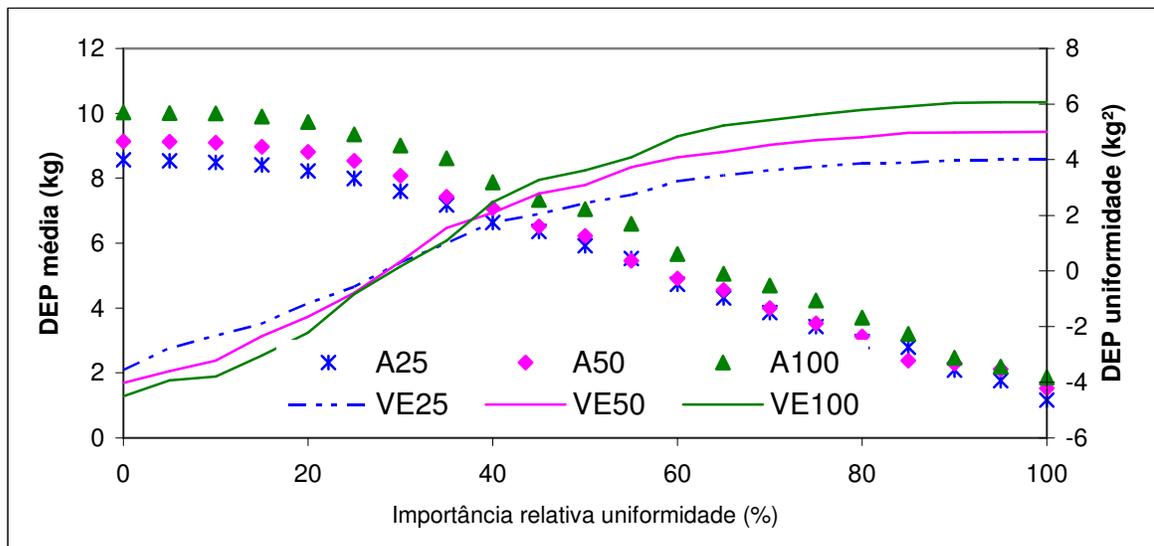


Figura 5. DEPs médias para tendência central (A25, A50 e A100) e uniformidade (VE25, VE50, VE100)\* do ganho de peso do nascimento à desmama de touros do decil 1 para o índice, mediante variação da importância relativa da uniformidade, considerando-se touros com, no mínimo, 25, 50 ou 100 filhos, respectivamente.

\*DEPs obtidas com base num modelo animal incluindo efeito de GC (LNE5)

Conforme observado no estudo de NEVES et al. (2008), utilizando outro estimador para cálculo da DEP para uniformidade de GND e outro banco de dados da raça Nelore, a consideração da uniformidade na composição do índice I implicaria em redução na resposta esperada sobre a média de GND, o que também estaria em concordância com a correlação estimada no presente estudo para as DEPs consideradas em I.

MULDER et al. (2008), ao simularem correlação desfavorável entre média e variância residual, numa situação de seleção dos animais com base num índice combinando estes dois critérios, também relataram ocorrência similar, sendo que ênfase substancial (cinco vezes maior) foi atribuída à uniformidade para que redução significativa na variância residual fosse observada, porém a custa de menor progresso genético da média da característica simulada.

No presente estudo, à medida que a seleção com base em I se tornasse mais intensa (seleção de touros com, no mínimo, 100 filhos) seria esperado maior progresso em termos de uniformidade, prevendo-se um incremento cerca de 50% maior na média das DEPs para uniformidade de GNDa, comparado à situação em que se consideraram touros

com, no mínimo, 25 filhos (A25 e VE25), nos cenários em que se atribuiu importância relativa acima de 45% para uniformidade.

Nos cenários em que seria esperada resposta sobre a uniformidade de GNDa (valores de  $u$  a partir de 30%), a média de DEPu aumentou de maneira considerável à medida que se atribuiu maior importância para uniformidade. Todavia, os ganhos médios esperados a cada incremento de 5 unidades percentuais em  $u$  praticamente se estabilizaram a partir de  $u = 80\%$ . A partir deste ponto, não seria esperado incremento significativo na uniformidade, ao passo que ainda haveria redução considerável no mérito para a média de GNDa.

Considerando-se seleção exclusiva para uniformidade ( $u=100\%$ ) e supondo que os touros fossem acasalados com vacas de mérito médio para este critério, seria esperada uma redução entre 14 e 17% no desvio-padrão residual de GNDa, em relação à média da geração anterior, dependendo do conjunto de dados considerado. Porém, esta situação implicaria em redução de mais de 80% na resposta para a média de GNDa, comparado a resposta esperada sob uma situação convencional ( $u=0$ ).

Ainda assim, possivelmente em virtude da baixa correlação entre estes dois critérios, seria possível progresso genético tanto na média quanto na uniformidade de GNDa por meio da seleção com base em  $I$ , sobretudo em situações em que a média de GNDa já estivesse mais próxima do ótimo.

Considerando uma situação de “penalização” no preço pago por animais com peso fora de um intervalo ótimo, seria possível exemplificar a vantagem econômica em se utilizar um touro melhorador para uniformidade da progênie. Comparando as DEPs de dois touros no arquivo de dados (touro A e touro B), com mérito similar para a média de GNDa (aproximadamente, +10,2 kg) e DEPs opostas para a variância residual de GNDa (+14,17kg<sup>2</sup> e -6,05 kg<sup>2</sup> para A e B, respectivamente), pode-se predizer, ainda que de modo simplista, a distribuição esperada do peso à desmama das progênies destes touros, supondo acasalamento aleatório com vacas de mérito médio para ambos os critérios e a influência dos mesmos efeitos ambientais sobre a progênie dos touros (Figura 6).

Assumindo-se um intervalo ótimo de 190-220 kg para peso a desmama e preço base igual a R\$580 por animal (com penalização de 10% no preço de animais fora desse intervalo), seria esperado um preço de venda, em média, 1,14% superior no caso da

progênie do touro B. Sendo assim, o benefício econômico do uso do touro B em detrimento do touro A seria tanto maior quanto maior o número de progênies, chegando a uma diferença de, aproximadamente, R\$ 1248,00, ao considerarem-se 200 filhos de cada touro.

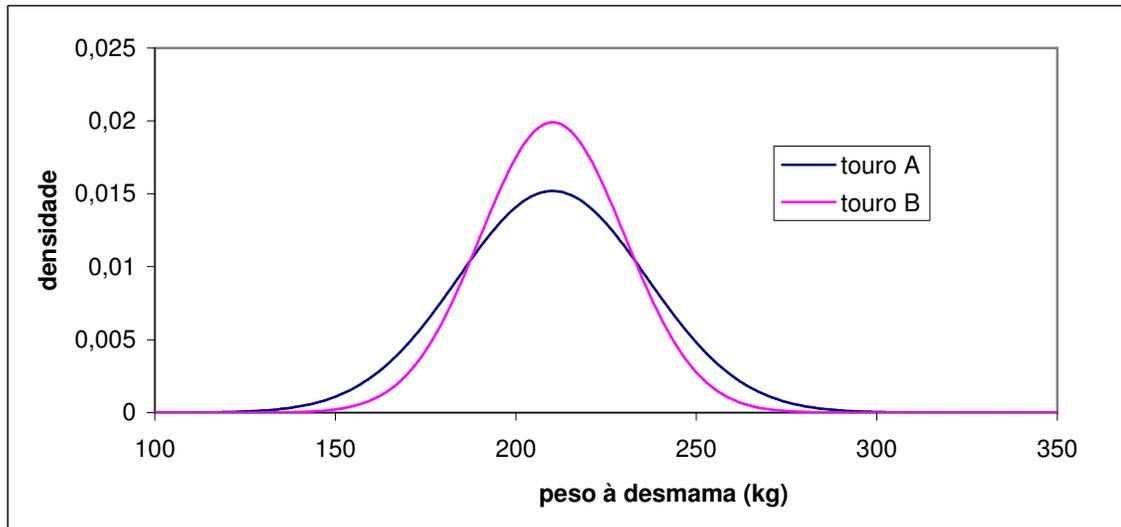


Figura 6. Distribuição esperada do peso à desmama da progênie de dois touros (A e B) com mesmo mérito para a média de GND e para peso ao nascer e mérito oposto para uniformidade de GND.

Apesar de se tratar de um exemplo teórico, tais resultados ilustram como a magnitude das diferenças observadas entre as DEPs dos touros para variância residual se traduziria em termos econômicos, tendo-se observado situação semelhante àquela relatada no estudo de IBAÑEZ-ESCRICHE et al. (2008). Estes autores simularam a aplicação de uma equação de lucro não-linear, contemplando faixas ótimas para peso de suínos ao abate e compararam dois reprodutores com mérito similar para a média e mérito oposto para variância residual. Neste caso, se ambos os reprodutores tivessem 500 filhos, a vantagem econômica para aquele com maior mérito para uniformidade (menor variância residual) chegaria a €2 por filho, totalizando um benefício econômico de €1000.

Analisando-se a tendências das médias dos EBVs para a média de GNDa (Figura 7), fica evidente o progresso genético para este critério, considerando-se o incremento no valor predito para efeito aditivo direto, observando-se uma amplitude de cerca de 11,2 kg no período observado. No caso do efeito materno sobre a média de GNDa, houve uma

tendência de ligeiro decréscimo, porém com a média deste valor ficando praticamente estável, a partir de 1990.

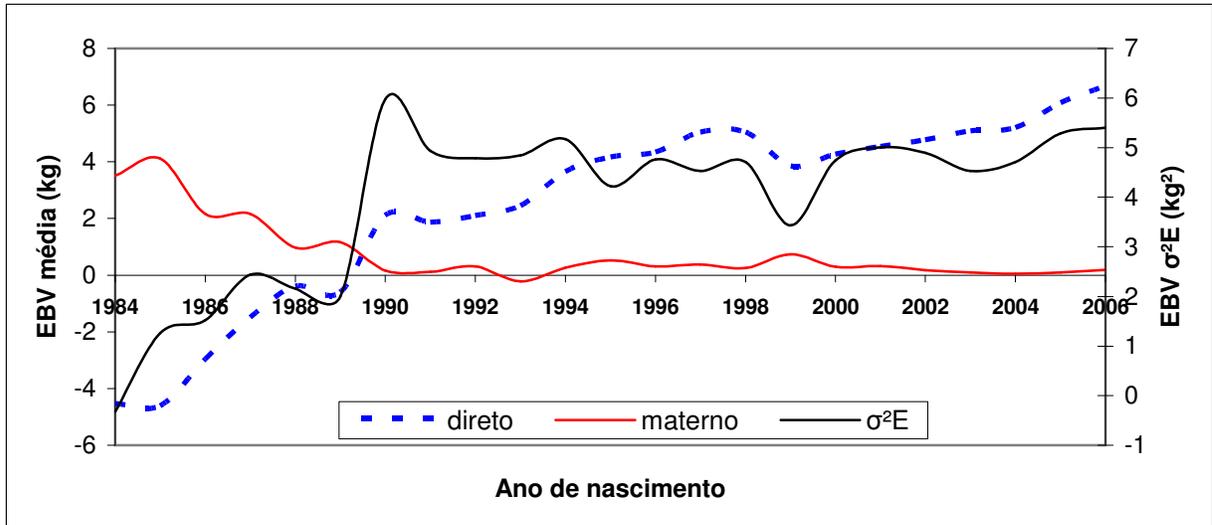


Figura 7. Tendências das médias das predições dos valores genéticos para a tendência central (direto e materno) e para variância residual ( $\sigma^2_E$ ) de GNDa.

As médias de EBV para a variância residual de GNDa se comportaram de modo bastante similar às médias de EBV para efeito direto, o que poderia estar associado à correlação positiva entre estes dois critérios (estimada neste estudo com base nas DEPs obtidas em cada passo das análises. Tal ocorrência poderia refletir uma resposta correlacionada desfavorável em termos de uniformidade da progênie, associada à seleção para incremento na média de GND.

Além da ação aditiva de genes atuando tanto sobre a média quanto sobre a variância residual de GND (pleiotropia), o incremento nos EBVs para  $\sigma^2_E$  ao longo do período analisado também pode se dever em alguma medida ao efeito descrito por HILL (2004). Segundo cálculos derivados por este autor, em populações sob intensa seleção por truncamento, há uma maior probabilidade de que genótipos mais variáveis sejam selecionados, no caso da existência de heterogeneidade genética sobre  $\sigma^2_E$ .

O estudo desta questão trata de uma situação bastante complexa, uma vez que os animais desta população foram selecionados com base em índices considerando diferentes critérios de seleção ao longo deste período, o que também pode ter influenciado este

resultado, bem como outros efeitos podem estar envolvidos neste fenômeno biológico (e.g. ação de genes de efeito maior), de modo que seria interessante que futuros estudos procurassem investigar este tipo de associação.

#### 5.4. CONCLUSÕES

Há evidências de efeitos aditivos e ambientais influenciando simultaneamente a variância residual de GND, bem como de que a seleção para menor variância residual implicaria em alguma redução na resposta sobre a média desta característica. Ainda assim, considerável resposta à seleção para uniformidade de GND poderia ser esperada no caso de seleção para menor variância residual, sobretudo no caso de seleção intensa e consideração de famílias de touro de grande tamanho.

#### 5.5. REFERÊNCIAS

- BOLDMAN, K.G.; KRIESE, L.A.; VAN VLECK, L. D.; VAN TASSEL, C. P.; KACHMAN, S. D. 1995. **A manual for use for MTDFREML**: a set of programs to obtain of variance and covariances. Lincoln: Department of Agriculture /ARS. 120p.
- BOLET, G.; GARREAU, H.; JOLY, T.; THEAU-CLEMENT, M.; FALIERES, J.; HURTAUD, J.; BODIN, L. Genetic homogenization of birth weight in rabbits: indirect selection response for uterine horn characteristics. **Livestock Science**, v. 111, pp. 28–32, 2007.
- GARREAU, H.; BOLET, G.; LARZUL, C.; ROBERT-GRANIÉ, C.; SALEIL, G.; SANCRISTÓBAL-GAUDY, M.; BODIN, L. Results of four generations of a canalizing selection for rabbit birth weight. **Livestock Science**, v. 119, pp. 55-62, 2008.
- GUTIERREZ, J. P.; NIETO, B.; PIQUERAS, P.; IBÁÑEZ, N.; SALGADO, C. Genetic parameters for canalization analysis of litter size and litter weight traits at birth in mice. **Genetics Selection Evolution**, v. 38, pp. 445–462, 2006.
- HILL, W. G. Heterogeneity of genetic and environmental variance of quantitative traits. **Journal of the Indian Society of Agricultural Statistics**, v. 57, pp. 49-63, 2004.
- IBÁÑEZ-ESCRICHE, N.; VARONA, L.; SORENSEN, D.; NOGUERA, J.L. A study of heterogeneity of environmental variance for slaughter weight in pigs. **Animal**, v. 2, pp.19-26, 2008.

- MULDER, H. A., HILL, W. G.; VEREIJKEN, A.; VEERKAMP, R. F. Estimation of genetic variation in residual variance in female and male broiler chickens. **Animal**, v. 3, pp. 1673-1680, 2009.
- MULDER, H. A.; BIJMA, P.; HILL, W. G. Prediction of breeding values and selection responses with genetic heterogeneity of environmental variance. **Genetics**, Bethesda, v.175, pp. 1895-1910, 2007.
- MULDER, H. A.; BIJMA, P.; HILL, W. G. Selection for uniformity in livestock by exploiting genetic heterogeneity of residual variance. **Genetics Selection Evolution**, v. 40, n. 1, pp. 37-59, 2008.
- NEVES, H. H. R. ; CARVALHEIRO, R. ; PENA, C. D. O. ; QUEIROZ, S. A. Incremento de uniformidade de produção em bovinos de corte por meio de seleção. In: VII SIMPÓSIO BRASILEIRO DE MELHORAMENTO ANIMAL, 2008, São Carlos, SP. Anais...São Carlos: SBMA, 2008.
- ROS, M.; SORENSEN, D.; WAAGEPETERSEN, R.; DUPONT-NIVET, M.; BONNET, J. C.; MALLARD, J. Evidence for genetic control of adult weight plasticity in the snail *Helix aspersa*. **Genetics**, v. 168, pp. 2089-2097, 2004.
- ROWE, S. J.; WHITE, I. M.S., AVENDAÑO, S.; HILL, W. G. Genetic heterogeneity of residual variance in broiler chickens. **Genetics Selection Evolution**, Les Ulis, n. 38, pp. 617-635, 2006
- WOLC, A. WHITE, I. M. S.; AVENDANO, S.; HILL, W. G. Genetic variability in residual variation of body weight and conformation scores in broiler chickens. **Poultry Science**, n.88, pp. 1156-1161, 2009.

## CAPÍTULO 6. IMPLICAÇÕES

Os resultados obtidos nos diferentes estudos permitem inferir que:

Há efeitos aditivos influenciando a variância residual de GND (VE) na população estudada, sendo que, ao se considerar famílias de grande tamanho, poder-se-iam obter predições acuradas do mérito genético dos touros e alguma resposta em termos de uniformidade de produção.

A não consideração de efeitos ambientais sobre VE no segundo passo das análises pode levar a superestimação da variância aditiva sobre a dispersão residual, bem como da resposta esperada à seleção.

A metodologia em dois passos investigada no presente estudo, segundo diferentes modelos, permitiu verificar a existência de efeitos aditivos e ambientais sobre a variância residual em bovinos de corte. Identificou-se superioridade do modelo que contempla efeitos aditivos sobre o logaritmo natural do quadrado do resíduo predito de cada observação fenotípica (modelo ALT), comparado ao uso de estimadores de dispersão dentro de touro (modelo DIS).

Há necessidade de mais estudos sobre a existência de heterogeneidade genética sobre a variância residual em bovinos de corte, visando identificar modelos mais adequados à estrutura de dados neste tipo de população, bem como permitindo eliminar eventuais efeitos de escala não considerados sob as abordagens estudadas.

A despeito da maior complexidade do ponto de vista computacional, também seria importante investigarem-se modelos contemplando a heterogeneidade genética sobre a variância residual com múltiplas características, inclusive permitindo avaliar como diferentes critérios de seleção (i.e. média e uniformidade destas características) estariam correlacionados. Este tipo de abordagem poderia tornar mais eficiente o processo de seleção em situações em que maior uniformidade de produto fosse valorizada, uma vez que a obtenção de resposta em termos de uniformidade de produto final provavelmente estaria condicionada a respostas em vários processos biológicos.

## APÊNDICE I – DERIVAÇÃO DO FATOR f, USADO PARA TRANSFORMAR AS VARIÂNCIAS ADITIVAS DOS ESTIMADORES DE ESCALA PARA A ESCALA DE VARIÂNCIA

A fórmula para a variância ( $\sigma^2$ ) da variância amostral de uma variável aleatória X (VAR, na notação deste trabalho), como derivada em Lynch & Walsh (1998) (Apêndice, p.814), para uma amostra de tamanho n é :

$$\sigma^2[\text{VAR}(X)] = [\mu_{4x} - \sigma^4(X)] / n$$

Se X é distribuída normalmente (como é pressuposto no caso dos resíduos das observações de GND), seu quarto momento central ( $\mu_{4x}$ ) se iguala a  $3\sigma^4(X)$ , levando a:

$$\sigma^2[\text{VAR}(X)] = 2\sigma^4(X) / n \quad (I)$$

A fórmula aproximada para a variância do desvio padrão amostral de X (SD) derivada em Lynch & Walsh (1998) (Apêndice, p.820) é:

$$\sigma^2[\text{SD}(X)] = (\mu_{4x} - \sigma^4(X)) / (4n\sigma^2(X))$$

O que leva à expressão:

$$\sigma^2[\text{SD}(X)] = 2\sigma^4(X) / (4n\sigma^2(X)) \quad (II)$$

Deste modo, a razão entre (I) e (II) pode ser calculada por:

$$f = \{ \sigma^2[\text{VAR}(X)] / \sigma^2[\text{SD}(x)] \} = 4 * \sigma^2(X).$$