

FELIPE GOMES DA SILVA

**AVALIAÇÃO GENÉTICA DA PRODUÇÃO DE LEITE NO DIA DO
CONTROLE DE CAPRINOS DE APTIDÃO LEITEIRA UTILIZANDO
MODELOS DE REGRESSÃO ALEATÓRIA MULTICARACTERÍSTICOS**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de Doctor Scientiae.

VIÇOSA
MINAS GERAIS – BRASIL
2014

**Ficha catalográfica preparada pela Biblioteca Central da Universidade
Federal de Viçosa - Câmpus Viçosa**

T

S586a
2014
Silva, Felipe Gomes da, 1986-
Avaliação genética da produção de leite no dia do controle
de caprinos de aptidão leiteira utilizando modelos de regressão
aleatório multivaracterísticos / Felipe Gomes da Silva. – Viçosa,
MG, 2014.

x, 76f. : il. : 29 cm.

Orientador: Robledo de Almeida Torres.
Tese (doutorado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Melhoramento genético. 2. Leite - Produção. 3. Modelos
de regressão aleatória. I. Universidade Federal de Viçosa.
Departamento de Zootecnia. Programa de Pós-graduação em
Genética e Melhoramento. II. Título.

CDD 22. ed. 636.30821

FELIPE GOMES DA SILVA

**AVALIAÇÃO GENÉTICA DA PRODUÇÃO DE LEITE NO DIA DO
CONTROLE DE CAPRINOS DE APTIDÃO LEITEIRA UTILIZANDO
MODELOS DE REGRESSÃO ALEATÓRIA MULTICARACTERÍSTICOS**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de Doctor Scientiae.

APROVADA: 12 de setembro de 2014.


Henrique Torres Ventura


Moysés Nascimento


Antônio Policarpo Souza Carneiro


Fabyano Fonseca e Silva
(Coorientador)


Robledo de Almeida Torres
(Orientador)

AGRADECIMENTOS

A Universidade Federal de Viçosa, principalmente ao programa de pós-graduação em Genética e Melhoramento, pela oportunidade de realização deste curso e ao departamento de Zootecnia fornecedor de estrutura para captação e análise dos dados.

A CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) e ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), pela concessão da bolsa de estudos, cada uma em uma fase do Doutorado. Junto a essas agências de fomento ao desenvolvimento científico e tecnológico, agradeço também à FAPEMG (Fundação de Amparo à Pesquisa do Estado de Minas Gerais), por ambas terem financiado projetos correlacionados e bolsas de iniciação científica para estudantes que colaboraram nesta pesquisa.

Ao meu orientador Professor Robledo de Almeida Torres, pelos ensinamentos, compreensão e cobrança, amizade e orientação, formação profissional e pessoal, etc. Ao Professor Fabyano Fonseca Silva, pela coorientação, participação das bancas de qualificação e defesa, críticas, sugestões, ensinamento e amizade. Aos Professores Antônio Policarpo de Souza Carneiro e Moysés Nascimento e ao Doutor Henrique Torres Ventura, pela participação na banca de qualificação e de defesa de tese e pelas críticas e sugestões que ajudaram a enriquecer minha formação.

Ao Professor Marcelo Teixeira Rodrigues pela concessão dos dados do setor de Caprinocultura da Universidade Federal de Viçosa, pela amizade, críticas e ensinamentos. E aos professores da UFV, aos quais agradeço através dos nomes principais, Ricardo Frederico Euclides, Paulo Sávio Lopes e Simone Elisa Facione Guimarães, pela grande contribuição que forneceram para minha formação profissional.

A Deus, pelas constantes graças alcançadas em minha vida. A minha esposa Lívia, exemplo de amor, caráter e dignidade. A minha mãe, Maria José, exemplo de compreensão. Ao meu Pai, João Oscar, exemplo de superação. A minha irmã, Elza, exemplo de senso de justiça. A minha irmã, Elisa, exemplo de solidariedade. A minha sobrinha, Filipa, exemplo de renovação. A minha vó Adalgisa, exemplo de força. Ao tio Robledo, exemplo de sabedoria. A tia Mirinha, exemplo de felicidade. Ao Robledinho, exemplo de planejamento. A Livinha, exemplo de esforço. ...

Aos amigos e colegas de trabalho que ajudaram na coleta dos dados e/ou análises: Ana Lúcia, Bianca, Delvan, Edson Costa, Elisa Maria, Fabiana, Fernanda, Fernando, Geraldo Iria, Gilberto Menezes, Giovani Costa, Luciano Silva, Luiz Brito, Hinayah

Oliveira, Joashileni, Lais Brito, Marjori Menezes, Nadson Souza, Letícia, Renata Veroneze e Vinícius Junqueira.

Aos amigos do melhoramento, pela amizade e troca de conhecimento: Alex, Ali Canaza, Aline, André Hidalgo, Ariane Tibúrcio, Bianca, Bruno Bastos, Camila, Carla Leite, Carlos, Carolina, Cristina Bonafé, Débora, Edson Costa, Geraldo, Hiago, Jeferson Correia, Joashisleni, José Lindemberg, Karol Nascimento, Katiene, Lorena Oliveira, Luana, Lucas Verardo, Luciano Silva, Marcela Barbieri, Marcos Lagrotta, Mariele Freitas, Mario Santana, Matilde, Nadson, Patricia, Rafael Bastos, Rodrigo Mota, Rodrigo Pacheco, Rodrigo Pereira, labtec & cia.

Aos colegas e amigos que me receberam muito bem no estágio na EMBRAPA Gado de Corte, os quais agradeço através dos pesquisadores Roberto Torres, Gilberto Menezes e Luiz Otávio e os estudantes, Mariela Marçal, Fabio, etc.

Aos colegas do Departamento de Zootecnia e Extensão Rural pela amizade, compreensão e colaboração. Obrigado aos professores Aline, Argon, Carlos Gondim, Carlos Cabral, Everton, Ferdinando, Heder, Ísis, Janessa, Joadil, Lisiane, Lucas, Luciano, Maria Fernanda, Marinaldo, Nelcino, Nicolau, Regina, Rosemary, Vanessa, Vânia, William. E também de outros departamentos, como os professores: André, Gerusa, Sandra, Sânia, Rosa, Janaína, Daniel, Nadia, Pedro, Emílio, etc. E aos alunos da UFMT, que sempre me surpreendem e com os quais muito aprendo e creço pessoalmente. Aos funcionários da FAMEVZ, Miguel, Denise, Monalisa, Cláudio, etc. pela colaboração nas atividades do dia dia.

Aos funcionários do capril, principalmente Antônio, Celso, José Maria, Ronaldo e Manoel, que foram muito importantes na execução da parte prática deste trabalho. Aos professores e funcionários do programa de Genética e Melhoramento por sempre terem me atendido muito bem e auxiliado nas duvidas e resolução de problemas e aos professores, funcionários e amigos do Departamento de Zootecnia, que é o segundo lar dos estudantes do melhoramento animal.

BIOGRAFIA

FELIPE GOMES DA SILVA, filho de João Oscar da Silva e Maria José Gomes da Silva, nasceu em 31 de janeiro de 1986, na cidade de Miradouro, estado de Minas Gerais, Brasil.

Em 2005, iniciou o curso de graduação em Agronomia na Universidade Federal de Viçosa, em Viçosa – MG, onde participou de estágios e foi bolsista de iniciação científica do CNPq e da FAPEMIG, na área de Melhoramento Genético Animal. Em 2008 participou do intercâmbio estudantil MARCA/MERCOSUL, cursando um semestre do curso de Agronomia na Universidad Mayor de San Andrés, La Paz – Bolivia. Em julho de 2009, obteve o diploma de Engenheiro Agrônomo pela Universidade Federal de Viçosa.

Em agosto de 2009, iniciou o curso de Mestrado em Genética e Melhoramento, na área de Melhoramento Animal, na mesma instituição. E em julho de 2011 submeteu-se a defesa e obteve o título de Magister Scientiae em Genética e Melhoramento.

Em agosto de 2011, iniciou o curso de Doutorado em Genética e Melhoramento, na área de Melhoramento Animal, na mesma instituição. Em novembro de 2012, foi aprovado em concurso público para Professor Assistente na Universidade Federal de Mato Grosso, Campus Cuiabá-MT, na área de Genética e Melhoramento dos Animais Domésticos. E em 12 de setembro, de 2014, submeteu-se aos exames finais de defesa da tese para obtenção do título de Doctor Scientiae em Genética e Melhoramento.

SUMÁRIO

| | |
|----------------------------------|------|
| RESUMO | viii |
| ABSTRACT | x |
| INTRODUÇÃO GERAL | 1 |
| REVISÃO DE LITERATURA | 4 |
| A - Caprinos | 4 |
| B - Avaliações Genéticas | 8 |
| BIBLIOGRAFIA CONSULTADA | 16 |
| CAPÍTULO 1 | 19 |
| Introdução | 20 |
| Material e Métodos | 21 |
| Resultados e Discussão | 24 |
| Conclusão | 31 |
| Referências Bibliográficas | 31 |
| CAPÍTULO 2 | 33 |
| Introdução | 34 |
| Material e Métodos | 35 |
| Resultados e Discussão | 37 |
| Conclusões | 42 |
| Referências Bibliográficas | 42 |
| CAPÍTULO 3 | 44 |
| Introdução | 45 |
| Caracterização | 46 |
| Exemplo Aplicado | 51 |
| Considerações finais | 55 |
| Referências Bibliográficas | 55 |
| CAPÍTULO 4 | 57 |
| Introdução | 58 |
| Material e Métodos | 60 |
| Resultados e Discussão | 62 |
| Conclusões | 65 |
| Referências Bibliográficas | 65 |

| | |
|---------------------------------|----|
| CAPÍTULO 5 | 67 |
| Introdução..... | 68 |
| Material e Métodos..... | 69 |
| Resultados e Discussão | 70 |
| Conclusões..... | 75 |
| Referências Bibliográficas..... | 75 |

RESUMO

SILVA, Felipe Gomes da, D.Sc., Universidade Federal de Viçosa, setembro de 2014. **Avaliação genética da produção de leite no dia do controle de caprinos de aptidão leiteira utilizando modelos de regressão aleatória multicaracterísticos.** Orientador: Robledo de Almeida Torres. Coorientador: Fabyano Fonseca e Silva.

O presente trabalho teve como objetivo principal realizar a avaliação genética da produção de leite no dia do controle utilizando modelos de regressão aleatória multicaracterísticos (MRAM). Os objetivos específicos foram: Estudar as divergências genéticas entre as quatro primeiras lactações de caprinos da raça Alpina, utilizando modelos de regressão aleatória unicaracterísticos (MRAU) para cada lactação; demonstrar a aplicação do software de recodificação de pedigrees alfanuméricos RENPED; verificar diferentes alternativas nas estruturas de ponto de truncamento, efeitos fixos e unidades de medidas da variável resposta produção de leite no dia do controle sobre a classificação dos animais avaliados geneticamente; verificar o melhor modelo de regressão aleatória para avaliar a produção de leite no dia do controle de caprinos da raça Saanen. Para tal, foram utilizados dados do rebanho de caprinos de aptidão leiteira da Universidade Federal de Viçosa. O software WOMBAT foi utilizado em todas as avaliações genéticas, o software CALC da LibreOffice foi utilizado em edições dos bancos de dados e o software RENPED foi utilizado nas recodificações de pedigrees e bancos de dados e algumas análises descritivas. Diferenças entre a estrutura genética de cada uma das lactações devem ser estudadas para delinear adequadamente estratégias de seleção em rebanhos de aptidão leiteira. Futuras análises testando diferentes modelos para a segunda, terceira e quarta lactação devem ser testados. Análises utilizando os MRAM foram viáveis, mas apresentam maior dificuldade de convergirem do que análises separadas de MRAU para cada lactação. MRAM e MRAU apresentam estimativas semelhantes para os componentes de variância ao longo da curva de lactação. MRAM podem trazer grandes vantagens em relação a MRAU por aproveitarem melhor os dados de ao menos duas lactações simultaneamente. O RENPED é uma alternativa para corrigir erros e recodificar dados alfa numéricos, assim como para viabilizar os arquivos de pedigree e dados para análises em diversos programas. Existe grande diferença na classificação dos animais quando o ponto de truncamento é muito baixo em relação a pontos de truncamento médios a altos para a lactação. Melhores resultados são obtidos em modelos que utilizem o efeito fixo classificatório ano-estação do controle leiteiro ao invés de ano-estação do parto. Melhores resultados também são obtidos através da construção destes efeitos fixos agrupando estações que comecem em um ano e terminem

em outro ao invés de apenas mesclar as variáveis ano e estação. A utilização da produção de leite no dia do controle em litros como tradicionalmente utilizado não atrapalha as avaliações genéticas através de problemas de erros numéricos, mas unidades de medidas muito grandes ou muito pequenas podem ocasionar este tipo de problema. O modelo mais parcimonioso para avaliação da PLDC para os caprinos da raça Saanen foi o que considerava regressão fixa de ordem 3, genética aditiva de ordem 2 e permanente individual de ordem 8, além de 5 classes de variância residual. Também concluiu-se que combinações de mesma ordem para as curvas genética aditivas e permanente individuais podem não encontrar o melhor modelo de acordo com os critérios de comparação tradicionais.

ABSTRACT

SILVA, Felipe Gomes da, D.Sc., Universidade Federal de Viçosa, September of 2014. **Study of Milk Yield on the Control Day of Alpine Goats Using Multitrait Random Regression Models.** Adviser: Robledo de Almeida Torres. Co-adviser: Fabyano Fonseca Silva.

The main objective of the present study was to conduct a genetic evaluation of test-day milk yield (TDMY) using multi-trait random regression models (MRRM). The specific objectives were: to study the genetic divergences between the first four lactations of Alpine goats using single-trait random regression models (SRRM) for each lactation; to demonstrate the application of the RENPED software to recode alphanumeric pedigrees; to determine different alternatives in the structures of truncation point, fixed effects and units of measurement of the TDMY response variable on the classification of animals evaluated genetically; and to determine the best random regression model to evaluate TDMY for Saanen goats. For this purpose, data from the dairy goat herd of Universidade Federal de Viçosa were used. The WOMBAT software was employed in all genetic evaluations; the CALC software from LibreOffice was used in datasets editions; and the RENPED software was used in the recoding of pedigree and datasets and a few descriptive analyses. Differences between the genetic structures of each one of the lactations should be studied to properly design selection strategies in dairy herds. Future analyses testing different models for the second, third and fourth lactations should be carried out. Analyses using the MRRM were feasible, but showed greater difficulty to converge than the separate SRRM analyses for each lactation. MRRM and SRRM show similar estimates for the variance components over the lactation curve. MRRM can provide greater advantages in relation to SRRM as they better utilize the data of at least two lactations simultaneously. RENPED is an alternative to correct errors and recode alphanumeric data as well as to make pedigree files and data available for analyses on various programs. There is a large difference in the classification of animals when the truncation point is too low in relation to medium and high truncation points for lactation. Better results are obtained in models that utilize the fixed year-season classificatory effect of the milk control through the construction of these fixed effects by clustering seasons that start in one year and end in another instead of only mixing the year and season variables. The use of TDMY in liters as traditionally adopted does not undermine the genetic evaluations with problems of numerical errors, but too large or too small units of measurement may generate this type of problem. The most parsimonious model to

evaluate TDMY for Saanen goats was the one which considered a fixed regression of order 3, additive genetic effect of order 2, permanent individual effect of order 8, in addition to five classes of residual variance. It is also concluded that same-order combinations for the additive genetic and permanent individual curves may not find the best model according to the traditional comparison criteria.

INTRODUÇÃO GERAL

A caprinocultura de leite tem crescente importância no cenário nacional, tendo alcançado 22.000 litros de produção diária em 2009 segundo dados do IBGE. Sendo que dentre as diferentes raças de caprinos de aptidão leiteira, as mais utilizadas no Brasil são as raças Saanen e a Alpina. Segundo Brito (2011), o leite de cabra é um alimento rico com características únicas. Além das vantagens nutricionais e terapêuticas, o leite de cabra possui alta qualidade proteica e grande conteúdo de minerais e vitaminas, melhor digestibilidade da fração lipídica e menor alergenicidade da fração proteica em relação ao leite de vaca.

Apesar de suas vantagens, seu mercado é muito aquém do esperado em relação ao leite de vaca. Possivelmente devido ao maior custo de produção ocasionado pelo déficit em pesquisas com caprinos em relação a pesquisas com bovinos, principalmente na área de melhoramento genético, além de um número muito reduzido de programas de melhoramento genético de cabras para produção de leite. A partir destas carências percebe-se a necessidade de se investir mais em pesquisas que aumentem a eficiência da produção de leite de cabras. Melhorando a nutrição, ambiência e principalmente o melhoramento genético para otimizar o sistema de produção de leite.

Os métodos de escolha de indivíduos geneticamente superiores de uma população estão em constante evolução. A seleção que começou de forma empírica, antes mesmo do surgimento da Genética como uma ciência, alcançou importantes resultados para a humanidade, principalmente no que diz respeito à domesticação das espécies. Entretanto, as técnicas de escolha empíricas possuem eficiência limitada dentro da biologia que coordena a herança dos caracteres.

Desta forma, diversas áreas da ciência têm sido combinadas para escolher os melhores indivíduos de uma população, equilibrando a razão do ganho genético econômico sobre a perda de variabilidade genética. A avaliação genética de caprinos de aptidão leiteira tem passado por diversos avanços, tais como avaliação de características visuais, avaliação da Produção de Leite Acumulada (PLA) através do modelo reprodutor, modelo animal, avaliação simultânea de mais de uma lactação através de modelos de Repetibilidade e modelos Multicaracterísticos. Mais recentemente, Modelos de Regressão Aleatória (MRA) tem sido utilizados com a vantagem de se avaliar a Produção de Leite no Dia do Controle (PLDC), ao invés da PLA.

De fato, MRA podem auxiliar a prática de seleção por aproveitarem melhor as informações da PLDC, serem menos sensível a dados perdidos, modelar as variações de

produção dentro da lactação, trabalhar com número maior de dados, etc. Também espera-se que os MRA permitam obter estimativas de parâmetros e predição de valores genéticos mais precisa e acurada, permitindo obter melhores ganhos genéticos. Isso decorreria do fato de ao invés de transformar as informações da PLDC em uma única informação, a produção de leite acumulada (PLA), os MRA permitem modelar todas as informações da PLDC ao longo dos dias em lactação (DEL).

Silva et al. (2013b) verificou os melhores Polinômios Ortogonais de Legendre para avaliação da PLDC em caprinos da raça Alpina realizando 980 análises. Silva et al. (2013c) demonstrou que utilizando o melhor MRA Unicaracterístico (MRAU), encontrados pelos autores anteriores para a primeira lactação, era possível aumentar o intervalo entre coletas da PLDC de 7 para 14 ou 21 dias, sem obter prejuízos nos resultados das avaliações genéticas, o que não ocorreu quando o intervalo entre controles era de 28 dias. Tal descoberta é fundamental para reduzir os custos com coletas de dados, visto que o aumento do custo de mão de obra confiável e qualificada é uma tendência mundial.

Apesar dos avanços, MRA trazem consigo alguns problemas, como por exemplo, como utilizar os resultados dos parâmetros da curva genética para praticar a seleção. Silva (2011c) resolve este problema utilizando o software RENPED que calcula o valor genético para cada dia em lactação e soma os valores genéticos de todos os dias da lactação para cada animal.

Um outro problema a ser considerado é que a avaliação da PLDC por meio de modelos de regressão aleatória não podia ser feita simultaneamente para mais de uma lactação e a seleção de animais apenas com base na primeira lactação pode acarretar em um tipo de animal que produz muito na primeira lactação, mas não se recupera adequadamente para ter um intervalo entre partos economicamente favorável nem a segunda e a terceira lactação, produtivas. Problemas facilmente resolvidos se forem realizadas uma análise para cada lactação.

Uma desvantagem das avaliações de lactações de ordem mais alta utilizando MRA é que o número de animais que apresentam medições nestas lactações reduz muito e o número de medições por animais também. Isso ocorre em função dos descartes e desgaste dos animais. Dessa forma lactações de ordem mais alta podem apresentar comportamentos de estimativas de componentes de variância distintas do esperado e menor acurácia dos valores genéticos preditos.

Acredita-se que tal problema possa ser amenizado utilizando o software WOMBAT (Meyer, 2007), o qual que traz recentemente a possibilidade de trabalhar com Modelos

de Regressão Aleatória Multicaracterísticos, aproveitando a estrutura de covariância entre parâmetros das diferentes lactações avaliadas simultaneamente. Dessa forma espera-se que a predição de valores genéticos para cada lactação seja feita de forma mais acurada, principalmente para as lactações de ordem mais elevada que se beneficiam mais da avaliação simultânea, por conterem menor volume de dados.

Partindo desta hipóteses que os Modelos de Regressão Aleatória Multicaracterísticos (MRAM) podem ser aplicados com maior eficiência do que MRA na avaliação da produção de leite no dia do controle de caprinos da raça Alpina e o aumento da acurácia seletiva e o tipo de animal mais completo que será possível selecionar através da utilização de MRAM compensará o aumento do custo computacional e da complexidade dos modelos a serem testados.

O objetivo principal a ser alcançado com este trabalho foi o de verificar se os MRAM podem ser aplicados adequadamente para avaliar a PLDC em caprinos da raça Alpina para as três primeiras lactações e avaliar a dificuldade de convergência e as complicações para se utilizar MRAM e fornecer retorno destas informações à comunidade científica.

Os objetivos secundários foram de estudar divergências entre os comportamentos da PLDC nas quatro primeiras lactações de caprinos da raça Alpina, verificar alterações nas estruturas de efeitos fixos a serem corrigidos em análises da PLDC, verificar divergências entre análises utilizando diferentes pontos de truncamento para o número de dias em lactação, verificar a ocorrência de erros numéricos nas análises, verificar o melhor MRAU utilizando Polinômios Ortogonais de Legendre, para avaliar a primeira lactação de caprinos da raça Alpina, demonstrar um exemplo aplicado do software de renumeração e recodificação de pedigrees (RENPED).

REVISÃO DE LITERATURA

A - Caprinos

Segundo Ribeiro (1997), os caprinos podem ser definidos como pequenos ruminantes, possuindo capacidade de aproveitamento de forragens e subprodutos. Segundo Ribeiro (1997) citado por Brito (2012), a cabra foi o primeiro animal capaz de produzir alimentos, cuja domesticação pelo homem ocorreu há cerca de dez mil anos, sendo os caprinos de mesma origem dos bovinos, com o tronco ancestral dos antílopes e a diferenciação ocorrendo no Plioceno.

De acordo com Silva (2011), as principais raças criadas no Brasil, de aptidão leiteira, são Alpina, Saanen e Toggenburg. Embora existam outras raças que podem ser encontradas em sistemas de produção de caprinos, tais como: Anglo Nubiana, Boer, Bhuj Brasileira, Marota, Canindé, Moxotó, Murciana e La Mancha. Mesmo que algumas destas tenham maior aptidão para produção de carne. De acordo com Ribeiro (1997), citado por Brito (2012), as raças domésticas atuais descendem provavelmente da *Capra aegagrus*, da Pérsia e Ásia Menor, *Capra falconeri*, do Himalaia, e da *Capra prisca*, da bacia do Mediterrâneo, sendo a cabra doméstica denominada *Capra hircus*.

A caprinocultura é uma atividade explorada em todos os continentes, estando presente em áreas que apresentam as mais diversas características edafoclimáticas (Brito, 2012). De acordo com Araújo (2004) o Brasil possui o décimo rebanho mundial de caprinos e o maior efetivo das Américas. Em sua revisão de tese Pires (2011) faz um importante resumo da história da chegada dos caprinos ao Brasil e evolução de populações destes animais.

No entanto, em alguns países, esta atividade apresenta reduzida expressão econômica, sendo, na maioria dos casos praticada de forma extensiva, adotando baixos níveis de tecnologia e, conseqüentemente, apresentando baixa produtividade e rentabilidade (Brito, 2012).

Cavalcanti e Silva (1988), também mostravam preocupação com este problema, ressaltando que 30 anos passados da importação de material genético, em quase totalidade do nordeste só se encontravam animais sem raça definida, o qual os autores atribuem a uma falta de atenção do setor público.

Talvez uma abordagem interessante fosse o fortalecimento de incentivo a formação de cooperativas para estabilizar a comercialização e fortalecer os produtores com assistência técnica de qualidade. Além de incentivar e fomentar associações de criadores

com avaliações genéticas realizadas por instituições brasileiras de ensino, pesquisa e extensão. Tais práticas aumentariam a qualidade genética dos rebanhos, aumentando os lucros com a produção, além da movimentação financeira ocasionada pela compra e venda de animais.

A incerteza da atividade e a falta dos incentivos públicos necessários também favorece a falta de cuidados e investimentos necessários. Mesmo práticas simples que podem evitar a transmissão de doenças não são frequentemente adotadas. Santa Rosa (1991) descreve algumas doenças que podem afetar rebanhos caprinos, tais como: Linfadenite Caseosa, Broncopneumonias, Pododermatite, Mamite, Clostridioses, Colibaciloses, Listerioses, Ceratoconjutivite, Micoplasmose, Chlamidiose, etc. Outras doenças comuns são Verminoses, Pododermite, Raiva, Aftosa, fotossensibilização, Brucelose, Kimeriose, Sarnas, Timpanismo, Actima, Mastites, etc.

Algumas doenças podem ser evitadas com medidas profiláticas simples, que na maioria das vezes não são adotadas. Outras, como a Artrite Encefalite Caprina (CAE), exigem um grande número de medidas para evitar o contágio entre os animais, sendo algumas mais simples e outras mais complicadas, por isso esta doença afeta até rebanhos com alto nível de tecnificação. Franke (1998) trás ensinamentos que permitem montar um plano de ação contra esta doença. Informações, mais recentes, sobre a CAE também podem ser encontradas em Silva et al. (2006).

Oliveira e Gonzales (1992) descrevem a transferência de embriões como uma alternativa para acelerar o melhoramento genético de caprinos. Mas a realidade dos criadores e sistemas de criação, custo e cuidados exigidos, e o problema da interação genótipo-ambiente, indicam que esta técnica só alcançaria resultados sustentáveis se fosse auxiliar a rebanhos com maior nível de tecnificação, exigindo mais uma vez políticas públicas coordenadas por instituições Brasileira de Pesquisa, Ensino e Extensão com profissionais competentes nas áreas de nutrição, bioclimatologia, sanidade, forragicultura e principalmente melhoramento genético.

Pimenta Filho e Sousa (1992) fazem uma crítica interessante às tentativas imediativas de melhoria do rebanho caprino no nordeste por meio de importações de materiais genéticos de outros países. Os autores relataram que este processo reflete em um progresso não duradouro e até prejuízo em alguns casos. Isso por efeito da interação genótipo-ambiente em que os caprinos exógenos não têm uma genética adaptada ao nordeste. Outro problema que pode ser percebido no texto é a respeito da falta de práticas adequadas como pesagens e controles leiteiros, registros de genealogia e seleção de animais mais produtivos. Tais práticas possuem resultados pequenos, mas duradouros, e

existe uma grande necessidade de selecionar os animais que já possuem genes de adaptabilidade para aumentar a produção de leite ou carne.

Segundo Nogueira Filho e Kasprzykowski (2006), citados por Brito (2012) 94% dos caprinos encontram-se em países em desenvolvimento, sendo uma importante fonte de alimento e renda para essas populações. Embora as pesquisas relacionadas à espécie caprina estejam aumentando em número e qualidade, também se nota, de forma desproporcional, o pequeno avanço de investimentos de indústrias e criações de grande porte.

Mesmo que de acordo com Costa (1984) o problema seja atribuído a um suposto preconceito em relação ao odor e peculiaridade do sabor. O que falta e muito a ser feito são investimentos na área de processamento de produtos de origem caprina, de forma a gerar produtos diferenciados. Também devem ser realizados muitos investimentos em marketing para reverter a imagem negativa que os produtos de origem caprina carregam até hoje.

Silva (1988) mostra dois gráficos de pesquisas realizadas pelo mesmo autor, um destes mostra uma pesquisa de opinião sobre qual o principal motivo do entrevistado que compra leite caprino o fazer. Dentre Apresentação, Antialérgico, Qualidade, Nutritivo, Sabor e Outros, 70% das escolhas recaem na menor alergenicidade deste produto. O que mostra que nesta época o leite de cabra era basicamente visto como um remédio para pessoas com problema de digestão de Alfa-Caseína-S1. O mesmo autor mostra em outro gráfico em que o consumo do queijo feito com leite de cabra, no estado da Bahia, cai drasticamente à medida que a renda da população aumenta. O que é um grande contrassenso quando comparamos atualmente os preços e padrões de queijos de cabras importados da França.

De acordo com Brito (2012), com base em dados do IBGE em 2010, a população caprina no Brasil é de aproximadamente 9,3 milhões de animais, com cerca de 90,8% deste efetivo, concentrado na região Nordeste. Dados do IBGE de 2011, não mostraram grande alteração deste efetivo, apresentando estimativa de 9,370 milhões de cabeças, sendo o estado da Bahia o de maior efetivo, 29,2% do total e a região Nordeste abrangendo 90,98% do efetivo, a Figura 1 serve de extrato dessa realidade.

Grandes barreiras a serem enfrentadas pela caprinocultura no Brasil são: Realizar marketing que mostre as qualidades diferenciadas dos produtos de origem caprina aos consumidores; Realizar políticas públicas que ajudem pequenos produtores a se agregarem em cooperativas e que criem associações de criadores que façam avaliações genéticas dos animais e forneçam assistência técnica; Vincular os órgãos anteriores a

instituições públicas de pesquisa, ensino e extensão, aproveitando as potencialidades dos profissionais destas instituições para auxiliar o crescimento sustentável da produção caprina.



Figura 1: Efetivo da população de caprinos no Brasil, de 1974 a 2009 e proporções do efetivo em cada uma das cinco regiões brasileiras de acordo com dados do IBGE 2011.

Dois raças de caprinos de aptidão leiteira são mais exploradas no setor de caprinocultura da Universidade Federal de Viçosa (UFV), as raças, Alpina e Saanen. Neste trabalho utilizamos animais com grau de sangue maior do que cinquenta por cento da raça Alpina nos capítulos de 1 até 4. O mesmo foi feito para a raça Saanen no capítulo 5.

A raça Alpina conhecida por muitos como “Parda Alpina” é originária dos Alpes Suíços e é encontrada tanto em regiões montanhosas como em regiões baixas. A raça Saanen é originária do Vale do Saanen, na Suíça. As cabras Alpinas possuem pelagem marrom com listra negra no dorso, enquanto as cabras Saanen possuem pelagem branca uniforme. No capril da UFV a média de produção destas cabras é de 2 kg de leite por dia. Mais detalhes podem ser encontrados em Brito (2012).

As raças Alpina e Saanen tem origem em regiões de clima temperado, a criação destes animais em climas tropicais embora haja trabalho de adaptação, têm resultados produtivos inferiores em relação à criação em clima temperado. Programas de melhoramento desenvolvidos em regiões de clima tropical podem ser úteis na redução desta distância, combinando os genes de maior produção (importação de material genético) com os de melhor adaptação ao clima local (processo de seleção adaptativa). Mais detalhes destas raças podem ser encontrados em Pereira (2004) e Ribeiro (1997).

Para o melhoramento genético de uma característica, é de fundamental importância conhecer o sistema de criação dos animais, para delinear estratégias para a seleção. O sistema de produção depende do sistema de criação, enquanto este depende da característica explorada, espécie, raça, clima, relevo, alimentação, etc. Em caprinos um importante fator a se considerar é a sazonalidade reprodutiva.

Silva (2011) traz no Apêndice III, de sua dissertação, uma discussão detalhada sobre os tipos de sistemas de criação de caprinos e o número de dias em lactação (DEL) a ser aferida a PLDC. Neste trabalho utilizamos uma DEL de 290 dias, pois permite realizar avaliações de maior abrangência.

B - Avaliações Genéticas

▪ Escolha dos efeitos fixos

Durante uma avaliação genética é de grande importância se definir efeitos fixos classificatórios e covariáveis a serem incluídas no modelo de forma a corrigir variações conhecidas. Tais variações, quando não corrigidas, podem reduzir a acurácia e a precisão da avaliação genética, porque a variação não atribuída ao efeito fixo classificatório ou covariável inflaciona o resíduo, além disso, uma parte dessa variação não controlada pode ser atribuída á variáveis aleatórias, prejudicando a predição dos valores genéticos e estimativa dos parâmetros genéticos.

Uma das formas de se escolher efeitos fixos previamente a uma avaliação genética é através do uso de uma análise de variância prévia. Santana Jr. et al. (2008) avaliaram a influência dos efeitos não genéticos de ano-estação de parto, tipo de parto, ordem do parto, mês do controle, dias em lactação e idade da cabra ao parto, sobre a produção de leite e porcentagem de componentes no leite. Apenas o fator, tipo de parto, não influenciou ($p < 0,05$) a produção de leite, indicando que os demais fatores deveriam ser incluídos em modelos de avaliação da produção de leite no dia do controle.

A inclusão de efeitos fixos deve ser testada, e não feita de forma equivocada, acrescentando-se todos os efeitos possíveis. Isso ocorre porque à medida que se aumenta o número de efeitos fixos, se perde em precisão nas subclasses combinando todos os efeitos fixos. A perda de precisão está ligada principalmente a redução do número de repetições dentro de uma subclasse, e tal perda pode ocasionar a atribuição de valores errados às classes de efeitos fixos, fazendo com que o que era para ser uma correção vire uma alteração prejudicial a análise (Silva, 2011).

Brito, et al. (2010) e Silva, et al. (2013a), avaliaram os Efeitos Fixos que afetam a PLA e PLDC respectivamente. O primeiro demonstra o procedimento adequado para se testar efeitos fixos a serem utilizados em Modelos Unicaracterísticos, Multicaracterísticos (**MM**) e de Repetibilidade, enquanto o segundo demonstra a forma adequada de se testar efeitos fixos a serem incluídos em MRA.

Silva, et al. (2013a) concluíram de forma interessante que independente dos resultados das análises dos efeitos fixos que devem ser incluídos em avaliações genéticas do rebanho utilizado em seu trabalho, todo e qualquer rebanho deve realizar o mesmo procedimento para avaliar quais efeitos fixos devem ser utilizados na avaliação genética. Ou seja, os efeitos fixos que influenciam uma população pode não ter efeito significativo em outra.

▪ **Modelos de Avaliação**

A seleção de animais para a produção de leite já foi calcada em considerar-se a informação individual da produção total, já foi praticada através do uso da informação de parentes, e passou por grande avanço ao introduzir o uso da metodologia de modelos mistos e selecionar os animais utilizando a melhor predição linear não-viesada (BLUP), em que o uso de todas as informações de parentesco permitiu aumentar a acurácia das estimativas dos valores genéticos dos animais.

Neste trabalho não realizaremos a revisão sobre modelos anteriores aos de Regressão Aleatória, tais como os modelos: Reprodutor e Animal Unicaracterísticos, de Repetibilidade e Multicaracterísticos para avaliar a PLA. Silva (2011) contém uma revisão mais detalhada sobre métodos de avaliação da Produção de Leite que pode ser utilizados para entender a evolução dos programas de melhoramento de animais de aptidão leiteira. Mais informações também podem ser encontradas em Silva et al. (2008).

Nos próximos tópicos abordaremos de forma sucinta os Modelos de Regressão Aleatória Unicaracterísticos (MRAU) utilizados por Silva et al. (2013a), Silva et al. (2013b) e Brito (2012) e os Modelos de Regressão Aleatória Multicaracterísticos (MRAM), que são o foco principal deste trabalho.

▪ **Modelos de Regressão Aleatória**

Henderson Jr. em 1982, propôs o uso de modelos de regressão aleatória para analisar dados longitudinais, contornando o problema de Super Parametrização das Análises Multicaracterísticas. Ao invés de estimar parâmetros e prever valores genéticos dos animais para cada característica, coloca-se a característica analisada em função do tempo, que pode ser, por exemplo, a idade dos indivíduos ou o número de dias em lactação. Então o que se faz é estimar um polinômio que explique a variação da característica em função do tempo, com uma parte fixa, média e geral para a população e uma parte aleatória permanente, contendo efeitos ambientais correlacionados, bem como desvios de interação intra-loco e entre locos distintos. E outra parte aleatória e genética

aditiva, com os desvios atribuídos aos efeitos genéticos aditivos dos animais. Sendo a parte aleatória desvios da produção dos animais em torno da curva fixa ajustada.

Desta forma os parâmetros estimados são dos regressores do polinômio ao invés de considerar a produção em cada idade, dia ou época como uma característica diferente. Tais modelos são muito convenientes, pois permitem prever valores genéticos e estimar componentes de variância em qualquer idade dentro do intervalo analisado. Espera-se melhores ganhos genéticos ao sairmos de uma avaliação utilizando o BLUP para uma única característica para MRAU, que utilizam diretamente todas as PLDC coletadas e levam em consideração a existência de uma estrutura de covariância entre estas medidas de PLDC (Silva, 2011).

Segundo Resende (2001), MRA podem ser considerados como modelos multivariados simplificados ou reduzidos que permitem a obtenção dos mesmos parâmetros de interesse que podem ser obtidos pelos modelos multivariados, mas com uma menor parametrização. Para se ter ideia, vamos supor que iremos analisar oito lactações sucessivas em caprinos. Por análise multicaracterísticas seriam necessárias a estimação de 72 parâmetros, 8 variâncias genéticas, 8 variâncias residuais, 28 covariâncias genéticas e 28 covariâncias de ambiente permanente. Utilizando MRA ao invés de modelos multicaracterísticos, teremos de estimar 7 parâmetros se considerarmos que um polinômio linear e 13 se considerarmos um polinômio quadrático. Se passarmos a considerar a produção de leite no dia do controle como uma característica, as vantagens de se utilizar MRAs se tornam conclusivamente mais adequadas (Silva, 2011).

MEYER (2004), trabalhando com dados simulados de crescimento de bovinos, encontrou acurácia, medidas pela correlação entre valores simulados e estimados, de 5,7 a 8,3% maior para MRA em relação a modelos multicaracterísticos. Segundo a autora a vantagem se deve a melhor modelagem das variâncias e outros parâmetros genéticos fornecidos no modelo de regressão aleatória. Entretanto, através da revisão feita até o presente momento é possível também imaginar que a menor parametrização dos MRA em relação a MM tenha reduzido os problemas de convergência dos algoritmos e devido à limitada capacidade dos computadores atuais para resolver análises superparametrizadas, tornando mais acurados os resultados das análises que utilizam MRA (Silva, 2011).

Outro fato importante a ser ressaltado, é que o uso de MRA não exige alterações no sistema de coleta, digitação e edição de dados, podendo-se usar as informações de bancos de dados já existentes para realizar avaliações genéticas. Por aproveitar melhor as informações, encontrar resultados mais acurados e as atividades de coleta de dados

representarem aumento nos custos de produção, esta metodologia pode reduzir o custo por litros de leite incrementado no valor genético dos animais.

O uso de MRA é muito desejável em situações em que as características estudadas são exemplos de dados longitudinais, e a medição em períodos específicos da vida do animal é difícil de ser cumprida. Isso ocorre muito nos sistemas de criação de bovinos e caprinos, tanto quando se analisa a produção de leite quanto se analisa o peso dos animais. Pois, embora haja mecanismos para sincronizar o parto das fêmeas, na prática não possui eficiência inquestionável e na verdade isto ainda não é praticado na maioria dos rebanhos, e adaptando-se a lógica de economia de escala, preparar a estrutura de medição para poucos animais mais frequentemente é mais custoso do que fazer menos medições em todos os animais de uma só vez.

A teoria de utilização de modelos de Regressão Aleatória pode ser resumida de forma superficial como modelos que avaliam os parâmetros de equações ao invés de avaliar cada dia de produção como uma característica. Além disso, apresentam uma curva fixa média para toda a população e duas curvas aleatórias específicas para cada indivíduo, uma que reflete em teoria o ambiente permanente que incide sobre o indivíduo e outra que dependerá dos laços genéticos e produções dos parentes, denominada curva genética aditiva. Revisões mais completas podem ser encontradas em Silva et al. (2008) e Silva et al. (2011).

Silva (2011) agrupou diversas vantagens da utilização de MRA na avaliação de dados longitudinais, citadas por vários autores, tais como Van der Werf e Schaeffer (1997), Kirkpatrick e Heckman (1989) e Albuquerque (2004). As principais são:

- 1) Permite que os dados sejam desbalanceados, ou seja, que os animais sejam pesados em épocas diferentes e que o número de medidas seja diferente, havendo soluções para os efeitos nas idades em que o animal não foi medido, embora estas soluções estejam restritas ao período de tempo entre as medidas tomadas mais cedo e mais tardiamente dentre todos os animais.

- 2) Quanto maior o número de medidas repetidas melhor é a vantagem de se utilizar MRA em detrimento de uma análise multicaracterística, devido à redução no número de parâmetros a serem estimados e valores genéticos a serem preditos.

- 3) Além de utilizar as informações de covariâncias entre as características para obter maior acurácia na estimativa de parâmetros genéticos e predição de valores genéticos, o faz de forma estruturada.

- 4) Permite obter estimativas para componentes de variância em todos os pontos dentro da trajetória, mesmo em períodos em que a característica não foi medida.

5) Evita erros ocasionados pela correção de características medidas em diferentes períodos para idades padrões.

6) Evita erros ocasionados pela estimação da produção total com base em produções parciais com correções para os períodos parciais.

7) Permite estudar de forma mais profunda o comportamento produtivo dos animais com respeito às características de interesse e planejar a seleção de forma mais adequada.

De acordo com Jamrozik et al. (1997), o uso de regressões para analisar a produção de leite no dia do controle não é recente. Entretanto, conforme relatado por Ptak e Schaeffer (1993), neste tipo de regressão o efeito de animal é utilizado apenas evidenciar a diferença entre os animais do rebanho. Schaeffer e Dekkers (1994) sugerem a possibilidade da utilização dos modelos de regressão aleatória em modelos lineares (Henderson, 1982) para analisar a PLDC. Dessa forma é possível combinar as vantagens de utilizar regressão e trabalhar com a PLDC com as vantagens de utilizar os efeitos aleatórios relativos apenas ao próprio indivíduo (parte de ambiente permanente e genética não transmissível) e relacionada ao parentesco (parte genética aditiva).

Utilizando MRAU, Silva et al. (2013b) demonstraram que intervalos entre controle de 14 e 21 dias também podem ser utilizados em detrimento de intervalos semanais no intuito de reduzir os custos com mão de obra gasta com a coleta, digitação e edição de dados. Entretanto intervalos entre controles de 28 dias apresentaram grandes alterações com relação aos indivíduos selecionados.

Dessa forma, esta é uma grande ferramenta a ser aplicada em programas de melhoramento praticados por empresas que queiram estabelecer grande competitividade, para programas de melhoramento de instituições de pesquisa, este aumento no intervalo pode não ser interessante, isto, pois a qualidade de modelos e intervalos entre controles, sempre serão alvos de novas pesquisas à medida que novos modelos teóricos precisam ser testados.

Mrode e Thompson (2005) ressaltam também a grande vantagem dos modelos de regressão aleatória na avaliação genética de animais de aptidão leiteira ao permitirem a seleção para maior persistência da lactação. Menezes et al. (2010) utilizou diferentes formulas para avaliar a persistência da lactação de caprinos da raça Saanen através de MRA, entretanto, os valores de herdabilidade encontrados para as medidas de persistência no rebanho em questão foram muito baixas, não sendo esperado grande ganho genético se a seleção fosse praticada com objetivo de aumentar a persistência.

Apesar dos modelos de regressão aleatória testados apresentarem grande vantagem sobre a estimação de valores genéticos para a PLDC ao invés da variável estimada PLA. Existe também uma grande desvantagem dos modelos propostos até agora que consiste em não considerar simultaneamente mais de uma lactação. Para contornar este problema, uma alternativa consiste na utilização de Modelos de Regressão Aleatória Multicaracterísticos, que permite estimar parâmetros para regressões de cada uma das lactações consideradas em uma mesma análise, aproveitando melhor as informações e estrutura de correlação entre as lactações.

▪ Modelos de Regressão Aleatória Multicaracterísticos

Kistemaker (1997) demonstrou, utilizando estudos de simulação, que os modelos de regressão aleatória avaliando a PLDC geravam resultados significativamente melhores que as análises tradicionais da PLA. Gerando incrementos de acurácia dos valores genéticos dos touros de 2 a 3% e das vacas de 6 a 8%.

Jamrozik et al. (1997), testaram a utilização de Modelos de Regressão Aleatória Multicaracterísticos (MRAM) para analisar a PLDC, a Produção de Gordura (PG), a Produção de Proteína (PP) e o Score de Células Somáticas (SCS), dentro de três lactações. Os autores chamaram esta tentativa de escolha lógica de teste frente às mudanças do sistema de coletas de dados produtivos do rebanho canadense. Espera-se dos MRAM melhores resultados frente ao aumento de características contínuas de interesse econômico a serem avaliadas em um rebanho.

Na conclusão apresentada por Jamrozik e seus colaboradores é possível perceber que a utilização de MRAM é muito interessante, pois permite aproveitar melhor diversas características longitudinais com grande volume de dados perdidos, o que é muito frequente quando rebanhos começam a ser avaliados pela PG, PP, etc. ao invés de apenas para a PLDC. Dessa forma se obtêm predições muito mais acuradas para os valores genéticos destas características que estão em ascensão.

Podemos estender o parágrafo anterior para vantagens esperadas na utilização de MRAM na avaliação de múltiplas lactações. Na prática o que ocorre é que muitas fêmeas são descartadas sem a avaliação de todas as lactações, muitas vezes o descarte nem se deve a baixa produção e sim a incidência de doenças não genéticas.

Konig et al. (2008) demonstraram uma interessante aplicação para os MRAM, ligada a avaliação de incidência de doenças no decorrer da lactação além da produção. Os autores explicaram que problemas de saúde nos animais podem aumentar a necessidade

de separação e tratamento especial, aumentar o custo com veterinários, e causar perdas econômicas na produção podendo até ser necessário o descarte do leite. Dois aspectos interessantes deste trabalho são utilização do modelo reprodutor, que pode facilitar a confecção de catálogos de touro de qualidade em grandes rebanhos, principalmente se a capacidade computacional for limitante. O segundo ponto interessante é que os dados de herdabilidade ao longo da curva indicam maior relação do componente genético aditivo com as doenças e produção, no final da curva de lactação. Podendo inclusive a persistência estar relacionada a resistência as doenças estudadas, o que indicaria que a seleção traria resultados em redução do custo e aumento da produção.

O programa WOMBAT, realiza análises de Modelos de Regressão Aleatória Multicaracterísticos, o exemplo 11 fornecido como suporte ao programa demonstra uma análise de um modelo de regressão aleatória bivariada.

▪ Critérios para escolha de um modelo

O critério básico para escolha de um modelo de regressão aleatória é que este traga resultados mais precisos e acurados sem aumentar significativamente a dificuldade para o pesquisador e a demanda computacional em relação aos modelos Unicaracterísticos e Multicaracterísticos propostos anteriormente.

Quando se tem um número razoável de dados, não se espera grandes diferenças nas estimativas de parâmetros genéticos entre as análises com cada um destes modelos. Espera-se que o MRA apresente estimativas mais acuradas, mas sua análise é mais complicada e computacionalmente mais difícil, desta forma grandes diferenças nas estimativas dos parâmetros genéticos entre as análises destes modelos podem indicar falhas no modelo, em rotinas de análises, na convergência dos algoritmos dentre outros problemas.

Na escolha do melhor modelo, contrasta-se o uso de diferentes equações para modelar a curva de produção de leite e consideram-se diferentes estruturas de variância residual. Entretanto são necessários critérios para escolher o melhor modelo para a situação analisada.

Já foi dito anteriormente que o sucesso de um análise utilizando um modelo de regressão aleatória, seja com objetivo de se estimar parâmetros genéticos ou avaliar geneticamente uma população, depende do polinômio ou função utilizada no modelo para descrever as variações da característica estudada em função do tempo. O modelo deve ser

adequado para explicar o fenômeno, permitindo separar corretamente os efeitos em fixos, aleatórios genéticos e aleatórios de ambiente permanente.

Entretanto, comparar modelos e julgar suas eficácias não é uma tarefa tão fácil e necessita seguir critérios para tornar a decisão correta. Na literatura são encontrados alguns critérios muito úteis para contrastar MRA, desde escolhas simples baseadas em menor variação da estimativa dos parâmetros genéticos ao longo da curva de lactação até a avaliação do tempo de análise cronometrado. Os critérios atualmente mais usados são: teste da razão de verossimilhança (LRT), critério do máximo do logaritmo da função de máxima verossimilhança (logL), critério da informação de Akaike (AIC) e critério da informação Bayesiano de Schwarz (BIC).

O número de parâmetros é também um critério, haja vista que modelos menos parametrizados são preferíveis devido ao menor tempo de análise e maior facilidade de convergência, espera-se também, menores erros de predições causados por ineficiência do conjunto software-hardwares utilizados na análise. Uma revisão mais completa sobre todos os critérios supracitados, para comparação de modelos, pode ser encontrada em Silva (2011).

BIBLIOGRAFIA CONSULTADA

- ALBUQUERQUE, L. G. **Regressão aleatória: Nova tecnologia pode melhorar a qualidade das avaliações genéticas**. V Simp. Soc. Bras. Melhoramento Animal. Pirassununga - SP, 2004.
- ARAÚJO, A.M. Paternidade e diversidade genética em Caprinos no Brasil por meio de Microssatélites de DNA. Dissertação. 2004.
- BRITO, L.F. Modelagem da produção de leite de cabras das raças Alpina e Saanen utilizando Regressão Aleatória. UFV, 2012. 82f. Dissertação (Mestrado), Universidade Federal de Viçosa.
- BRITO, L.F.; SILVA, F.G.; MELO, A.L.P.; CAETANO, G.C.; TORRES, R.A.; RODRIGUES, M.T.; MENEZES, G.R.O. Genetic and environmental factors that influence production and quality of milk of Alpine and Saanen goats. **Genetics and Molecular Research**. 3794-3802p. 2010.
- CAVALCANTI, G.C.; SILVA, R.C. **Aspectos da caprino-ovinocultura na região nordeste: Tecnologia, produção e comercialização**. Recife, PE. 36p. 1988.
- COSTA, J.P.C. Observações sobre a Caprino/Ovinocultura. Editora: EMATER-SE, Aracajú, SE. 57p. 1984.
- FRANKE, C.R. Controle Sanitário da Atrite-Encefalite Caprina (C.A.E.). Editora EDUFBA, Salvador, Bahia. 70p. 1998.
- HENDERSON JR., C. R. Analysis of covariance in the mixed model: higher-level, Nonhomogeneous and random regression. **Biometrics**, v.38, p.633-40,1982.
- IBGE - Instituto Brasileiro de Geografia e Estatística. Endereço virtual: www.ibge.gov.br. Acesso em 10 de janeiro de 2014.
- JAMROZIK, J.; SCHAEFFER, L. R.; LIU, Z. & JANSEN, G. 1997. Multiple trait random regression test day model for production traits. Vienna, Austria, Aug. 28-29, 1997. Interbull Bulletin, No. 16, 43.
- KIRKPATRICK, M.; HECKMAN, N. A quantitative genetic model for growth and other infinite dimensional characters. **Journal of Mathematical Biology**, v.27, p.429-450, 1989.
- KISTEMAKER, G. J. 1997. PhD Thesis. University of Guelph. 1997.
- KÖNIG, S., ET AL. "Application of a multivariate random regression sire model to estimate genetic parameters among milk yield, somatic cell count and laminitis." Interbull Bulletin 38 (2008): 83.

- MENEZES, G.R.O.; TORRES, R.A.; SARMENTO, J.L.R.; RODRIGUÊS, M.T.; MELO, A.L.P.; SILVA, F.G.; BRITO, L.F. Avaliação de medidas da persistência da lactação de cabras da raça Saanen sob modelo de regressão aleatória. **Revista Brasileira de Zootecnia**. v.39, n.8, p.1691-1698, 2010.
- MEYER, K. WOMBAT – A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). *Journal Zhejiang University Science B*, v. 8 n. 11, p. 815-821, 2007.
- MEYER, K. Scope for a random regression model in genetic evaluation of beef cattle for growth. **Livestock Production Science**, v.86, p.69-83, 2004.
- MRODE, R.A.; Thompson, R. **Linear Models for the prediction of Animal Breeding Values. Second Edition**. CABI Publishing. 344p. 2005.
- NOGUEIRA FILHO, A., KASPRZYKOWSKI, J. W. A., **O agronegócio da caprino ovinocultura no Nordeste Brasileiro**. Série Documentos do ETENE, n. 09, 56p. Fortaleza: Banco do Nordeste do Brasil, 2006.
- OLIVEIRA, V.S.; GONZALES, C.I.M. Transferência de Embriões em Caprinos. Anais do I Simpósio Nordestino sobre Caprinos e Ovinos Deslanados. Editora APACCO, Taperoá, PB. 1992.
- PEREIRA, J. C. C. **Melhoramento genético aplicado à produção animal**. 4ª ed. Belo Horizonte: Editora FEPMVZ 2004. 609 p.
- PIMENTA FILHO, E.C.; SOUSA, W.H. Bases para o melhoramento genético de caprinos leiteiros. Anais do I Simpósio Nordestino sobre Caprinos e Ovinos Deslanados. Editora APACCO, Taperoá, PB. 1992.
- PIRES, L.C. Diversidade Genética de Caprinos das Repúblicas de Cabo Verde e do Brasil. Viçosa, 2011. 123p. Tese (Doutorado), Universidade Federal de Viçosa.
- PTAK, E. and SCHAEFFER, L.R. 1993. Use of test day yields for genetic evaluation of dairy sires and cows. **Livestock Prod. Sci.**, 34(1-2):23-34.
- REZENDE, M. D. V.; REZENDE, G. D. S. P.; FERNANDES, J. S. C. Regressão aleatória e funções de covariância na análise de medidas repetidas. **Revista Mat.** São Paulo, 19, p.21-40, 2001.
- RIBEIRO, S. D. de A.; **Criação Racional de Caprinos**. São Paulo: Nobel, 1997. 318 p. ISBN: 85-213-0972-4.
- SANTANA JÚNIOR, M.L.; MELO, A.L.P.; PEREIRA, R.J.; SILVA, F.G.; et al. (2008). **Influência de Fatores Ambientais na Porcentagem dos Componentes do Leite de Caprinos da Raça Alpina**. In: VII Simpósio Brasileiro de Melhoramento Animal, São Carlos-SP, 2008.

- SANTA ROSA. Anais da 28^o Reunião Anual da Sociedade Brasileira de Zootecnia. João Pessoa, PB. 557p. 1991.
- SCHAEFFER, L. R., AND J. C. M DEKKERS. 1994. Random regressions in animal models for test-day production in dairy cattle. In: Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph, Canada. 18:443–446.
- SILVA, E.R.V.; MENEZES, A.T. and OLIVEIRA FILHO, J.P. (2006). Artrite Encefalite Caprina. Revista Científica Eletrônica de Medicina Veterinária. Publicação Científica da Faculdade de Medicina Veterinária e Zootecnia de Garça/FAMED. Ano III, Número 6, Janeiro de 2006.
- SILVA, F. G. Estudo da produção de leite de caprinos utilizando modelos de regressão aleatória. Viçosa, UFV, 2011. 143p. Dissertação (Mestrado), Universidade Federal de Viçosa.
- SILVA, F. G.; BRITO, L. F.; TORRES, R. A.; RIBEIRO JÚNIOR, J.I.; OLIVEIRA, H. R.; CAETANO, G.C.; RODRIGUES, M. T. Factors that influence the test day milk yield and composition. *Genetics and Molecular Research*. V.12, p.1522 - 1532, 2013a.
- SILVA, F. G.; TORRES, R.A.; BRITO, L. F.; SILVA, L. P.; MENEZES, G.R.O.; Brito, L. C.; RODRIGUES, M.T.; EUCLYDES, R. F. Genetic evaluation of Alpine goats using different milk control intervals. **Genetics and Molecular Research**. v.12, p.2455 - 2464, 2013b.
- SILVA, F.G.; TORRES, R.A.; BRITO, L.F.; EUCLYDES, R. F.; MELO, A. L. P.; SOUZA, N. O.; RIBEIRO, Jr. J.I. and RODRIGUES, M.T. Random regression models using Legendre orthogonal polynomials to evaluate the milk production of Alpine goats. **Genetics and Molecular Research**. v.12(4), p.6502-6511, 2013c.
- SILVA, M. A.; THIÉBAUT, J. T. L.; VALENTE, B. D.; TORRES, R. A.; FARIA, F. J. C. Modelos lineares Aplicados ao Melhoramento Genético Animal. Belo Horizonte: FEPMVZ Editora, 2008. 1^aed. 378p.
- VAN DER WERF, J.; SCHAEFFER. **Random Regression in Animal Breeding**. Course Notes, Ontario: University of Guelph, 1997, 70p.

CAPÍTULO 1

Avaliação genética da produção de leite de caprinos Alpinos para as quatro primeiras lactações utilizando modelos de regressão aleatória unicaracterístico

Resumo: A avaliação genética de caprinos de aptidão leiteira tem passado por diversos avanços. Modelos de regressão aleatória tem sido utilizados com a vantagem de se avaliar a produção de leite no dia do controle, ao invés da produção de leite acumulada, obtendo estimativas mais acuradas dos valores genéticos dos animais. Todavia cuidado deve ser tomado pois a avaliação dos melhores animais apenas com informações da primeira lactação pode não implicar na identificação dos melhores do rebanho do ponto de vista econômico. Outros fatores devem ser levados em consideração, principalmente as demais lactações. Objetivamos com este trabalho estudar a divergência genética entre as quatro primeiras lactações de caprinos da raça Alpina. O programa RENPED foi utilizado para realizar a estatística descritiva, a verificação de erros no pedigree, a recodificação dos dados e as correlações de Pearson e Spearman entre os valores genéticos. O programa WOMBAT foi utilizado para estimar componentes de variância e obter os valores genéticos dos animais avaliados. O programa CALC da LibreOffice foi utilizado para calcular a porcentagem de coincidência entre a ordem classificatória dos animais e as porcentagens de animais mantidos em comum na avaliação de cada lactação. Os resultados mostraram que a seleção apenas pela primeira lactação em rebanhos pequenos e com baixo grau de tecnificação pode ser utilizada como uma medida paliativa frente a dificuldade de se avaliar todas as lactações. Mas a seleção de reprodutores e confecção de catálogos não pode ser feita apenas considerando a primeira lactação, pois os resultados demonstram inversões na classificação dos melhores reprodutores quando outras lactações são analisadas.

Palavras-chave: descarte em comum, Legendre, valor genético.

Introdução

Os métodos de identificação de indivíduos geneticamente superiores em uma população estão em constante mudança ou processo de evolução. A seleção que começou de forma empírica, antes mesmo do surgimento da Genética como uma ciência, alcançou importantes resultados para a humanidade, principalmente no que diz respeito à domesticação das espécies.

As técnicas de escolha empíricas possuem eficiência limitada dentro da biologia que coordena a herança dos caracteres, desta forma, diversas áreas da ciência têm sido combinadas para escolher os melhores indivíduos de uma população, maximizando o ganho genético econômico com pouca perda de variabilidade genética.

A avaliação genética de caprinos de aptidão leiteira tem passado por diversos avanços, os quais nós podemos enumerar evolutivamente, como: avaliação empírica, avaliação de características visuais correlacionadas à produção de leite, avaliação da Produção de Leite Acumulada (PLA) através do modelo reprodutor, modelo animal, avaliação simultânea de mais de uma lactação através de modelos de Repetibilidade e modelos Multicaracterísticos. Mais recentemente, Modelos de Regressão Aleatória (MRA) tem sido utilizados com a vantagem de utilizar a Produção de Leite no Dia do Controle (PLDC), ao invés da PLA que é uma medida estimada utilizando valores de PLDC. Isso implica não só em um maior volume de dados, como também maior nível de detalhamento do fenótipo, o que pode permite obter estimativas dos valores genéticos mais acuradas e coerentes à biologia dos animais.

Brito et al. (2010) e Silva et al. (2013a) avaliaram os efeitos fixos que afetam a PLA e PLDC, respectivamente. O primeiro demonstra o procedimento adequado para se testar efeitos fixos a serem utilizados em Modelos Unicaracterísticos, Multicaracterísticos e de Repetibilidade, enquanto o segundo demonstra a forma adequada de se testar efeitos fixos a serem incluídos em MRA.

Embora muitos trabalhos testando MRA em animais de aptidão leiteira optem por funções paramétricas de Wilkink (FPW) em detrimento dos polinômios ortogonais de Legendre (POL), Menezes et al. (2010), trabalhando com caprinos do rebanho da UFV, obtiveram um POL como melhor modelo, superando as FPW testadas pelos autores. O melhor modelo encontrado pelos autores foi o POL que continha ordem quatro para variância genética aditiva e seis para os efeitos permanentes individuais.

Silva (2011) avaliando 980 MRA utilizando POL, demonstrou que os melhores MRA podem não ser os tradicionalmente testados. O modelo mais parcimonioso com

resultados iguais ao modelo completo, pelo teste da razão de verossimilhança (TRV) foi o que continha ordens quatro, dois e sete para as regressões fixa, genética aditiva e permanente individual respectivamente, além de cinco classes de heterogeneidade de variância.

Brito (2012) avaliou MRA utilizando funções do tipo B-Spline equidistantes, que trazem melhorias nas predições quando modelos muito parametrizados são exigidos na explicação do comportamento de dados longitudinais. Este tipo de modelo permite em muitos casos reduzir problemas de multicolinearidade nas matrizes de componentes de variância. Entretanto, o autor não encontrou nenhum modelo que superasse o melhor POL encontrado por Silva (2011), de acordo com os critérios AIC e BIC, para a mesma população. A eficiência de funções do tipo B-Spline não equidistantes podem ser melhores que os modelos testados.

Grandes melhorias foram alcançadas utilizando MRA, entretanto estes trazem consigo dois problemas. Um consiste em como utilizar os resultados dos parâmetros da curva para selecionar os animais e outro consiste na avaliação apenas de uma lactação por vez, o que não é adequado, pois os genes que influenciam cada lactação podem não ser os mesmos.

Silva (2011) sugere para o primeiro problema integrar a área abaixo da curva genética de cada animal, utilizando o software RENPED, o autor ainda apresenta o manual de utilização e o código fonte do programa. Lidauer et al. (2000) descrevem uma solução matricial para este problema envolvendo uma matriz de autofunções.

O segundo problema poderia ser contornado utilizando um índice que seria a soma dos valores genéticos acumulados encontrados para cada uma das lactações. Desta forma os animais de maiores valores resultantes, seriam aqueles que teriam a melhor genética para produção de leite em toda a vida produtiva do animal.

Portanto, objetivou-se com este trabalho estudar a divergência genética entre as quatro primeiras lactações de caprinos da raça Alpina utilizando modelos de regressão aleatória unicaracterísticos.

Material e Métodos

Foram utilizados dados da PLDC de cabras do setor de caprinocultura da Universidade Federal de Viçosa (Viçosa, MG). Os animais foram criados em sistema de estabulação livre e receberam alimentação à base de silagem de milho e feno, como

volumoso, e mistura concentrada, fornecida conforme a exigência nutricional dos animais.

Os controles leiteiros foram realizados semanalmente em ordenha mecânica, realizada duas vezes ao dia. A coleta na parte da manhã teve início às seis horas e a coleta na parte da tarde teve início às quatorze horas. A PLDC representa a soma do leite coletado nestes dois períodos.

Após a análise crítica dos dados restaram 50.405 registros de PLDC das quatro primeiras lactações de 851 cabras da raça Alpina no arquivo completo. Deste foram gerados quatro novos arquivos a serem analisados, cada um contendo informações de uma das quatro primeiras lactações.

As análises descritivas para os valores amostrais da PLDC foram obtidos utilizando o programa RENPED (Silva, 2011). O mesmo programa foi utilizado para corrigir erros de pedigree e recodificar os dados alfanuméricos, procedimentos que melhoram e viabilizam a análise dos dados, respectivamente.

O programa WOMBAT (Meyer, 2007), versão atualizada de setembro de 2013, foi utilizado em todas as análises genéticas, fornecendo as estimativas dos parâmetros e predições dos valores genéticos para cada animal avaliado.

Os modelos utilizados nas avaliações da PLDC continham combinações de ajuste para os polinômios das curvas fixa, genética aleatória e de ambiente permanente de ordem 4, 2 e 7, além do número de classes de variâncias residuais igual a 5. Seguindo o modelo recomendado por Silva et al. (2013b), para o mesmo rebanho. O modelo animal unicaracterístico de regressão aleatória utilizado em cada uma das quatro lactações é descrito de forma generalizada como:

$$y_{ij} = EF_i + \sum_{m=0}^{4-1} b_m \phi_m(t_{ij}) + \sum_{m=0}^{2-1} \alpha_{im} \phi_m(t_{ij}) + \sum_{m=0}^{7-1} \gamma_{im} \phi_m(t_{ij}) + \varepsilon_{ij}$$

em que: y_{ij} corresponde à produção de leite no controle j da cabra i ; EF_i consiste no conjunto de efeitos fixos, sendo a parte classificatória constituída da média, do grupamento genético (Alpina > 90% = 1; 80% < Alpina ≤ 90% = 2; 70% < Alpina ≤ 80% = 3; 60% < Alpina ≤ 70% = 4; 50% < Alpina ≤ 60% = 5), do ano-estação e tipo de parto. A idade da cabra ao parto foi incluída como covariável, com efeitos linear e quadrático; b_m é o coeficiente de regressão m da PLDC sobre o polinômio de Legendre para modelar a curva média da população; α_{im} e γ_{im} são os m coeficientes de regressão genéticos

aditivos e de ambiente permanente, respectivamente, para a cabra i ; 4-1, 2-1 e 7-1 são os graus dos polinômios ortogonais de Legendre; t_{ij} é a variável controle, semana de lactação da cabra i padronizada para o intervalo de -1 a 1 conforme descrito por Kirkpatrick et al. (1990); $\phi_m(t_{ij})$ é a função polinomial de Legendre para o parâmetro m avaliada para a idade t_{ij} ; e ε_{ij} consiste nos efeitos específicos à cada observação não explicados pelos fatores de correção nem pelas regressões do modelo.

Neste trabalho utilizamos os mesmos efeitos fixos encontrados em Silva et al. (2013a), para avaliar a PLDC, por se tratar de dados do mesmo rebanho. Em rebanhos ou características distintas, os autores recomendam a realização de novos testes.

Para estudar e entender divergências genéticas entre diferentes lactações foram realizadas todas as análises utilizando o mesmo modelo, caso contrário poderia haver confundimento entre a estrutura e o efeito de modelagem.

Critérios de comparação de modelos tais como logaritmo da função de verossimilhança restrita (Log L) e critério da informação de Akaike (AIC, por Akaike, 1973) não podem ser utilizados para comparar diferentes bancos de dados, pois foram propostos para comparar diferentes modelos sobre o mesmo conjunto de dados. Além disso, a utilização dos mesmos para contrastar diferentes lactações traria confundimento de efeitos de amostragem e volume de dados, não servindo para indicar divergências ou semelhanças nas estruturas genéticas que coordenam a produção de leite em cada lactação

De acordo com Misztal (2008), em muitos modelos, a convergência utilizando o algoritmo AI-REML (Average Information) é atingida em apenas algumas rodadas, no entanto, valores de partida de baixa qualidade, podem inviabilizar a convergência, ou encontrar estimativas incoerentes, inclusive fora do espaço paramétrico. De fato, temos percebido que o algoritmo AI-REML é muito veloz e apresentam estimativas admissíveis. Entretanto, quando os valores de partida são de baixa qualidade, torna-se comum que as estimativas saiam do espaço paramétrico e não alcancem convergência. Neste caso sugere-se utilizar outros algoritmos mais lentos, como o PX-EM, para obter valores de partida melhores para repetir as análises com o AI-REML, se no final das análises a conversão por esse algoritmo for alcançada os resultados são aceitos, caso contrário, considera-se que os resultados de outros algoritmos foram ruins, pois não serviram nem como bons valores de partida para o AI-REML.

Os resultados dos valores genéticos para os parâmetros do modelo aditivo foram integrados, utilizando o RENPED, para se obter o valor genético acumulado de cada

animal em cada lactação. O mesmo programa também foi usado para calcular as correlações de Pearson e Spearman entre os valores genéticos para cada lactação.

A função PROCV do programa Calc, da LibreOffice foi utilizado para calcular a porcentagem de coincidência da ordem classificatória dos animais e a porcentagem de animais mantidos em comum quando são descartados 20, 30 e 40% dos animais. Estes testes permitem verificar se os indivíduos selecionados na primeira lactação também são os melhores nas demais. Este mesmo programa também foi utilizado para calcular a correlação genética entre os diferentes dias em lactação (6 a 290) para cada lactação, utilizando processos matriciais.

Resultados e Discussão

A Figura 1 apresenta a flutuação dos componentes de variância (fenotípico, aditivo, permanente e residual) do modelo estudado em função do número de dias em lactação, para as quatro primeiras lactações.

Os resultados encontrados para as variâncias genéticas aditivas, efeitos permanentes e residuais não diferiram muito em comportamento, quando as quatro lactações são contrastadas. Embora a terceira lactação fuja um pouco ao comportamento das demais, principalmente por apresentar menor incremento da variância genética aditiva no final da curva, esperando-se um pouco de divergência quanto aos genes associados à persistência desta em relação às demais lactações. Isto pode estar também relacionado a alterações de manejo, alimentação ou exigências que possam requerer genes de adaptabilidade distintos nesta lactação em relação às demais ou mesmo a algum problema com as estimativas encontradas. Já para as lactações 1, 2 e 4, espera-se que os mesmos genes de produção, adaptabilidade, resistência a doenças estejam se manifestando de forma semelhante.

Em modelos de regressão aleatória, a variância fenotípica é ajustada no modelo pelas equações de regressão genética aditiva e de efeitos permanentes individuais. A grande diferença entre as duas equações é que a genética aditiva utiliza as informações de parentesco enquanto a de efeitos permanentes utiliza apenas as informações do mesmo animal sob uma estrutura de covariância. A regressão permanente é ainda por vezes denominada ambiente permanente, podendo por vezes estar equivocada, pois esta equação busca ajustar também efeitos genéticos não transmissíveis como a maior parte dos desvios de interação intra e entre loco.

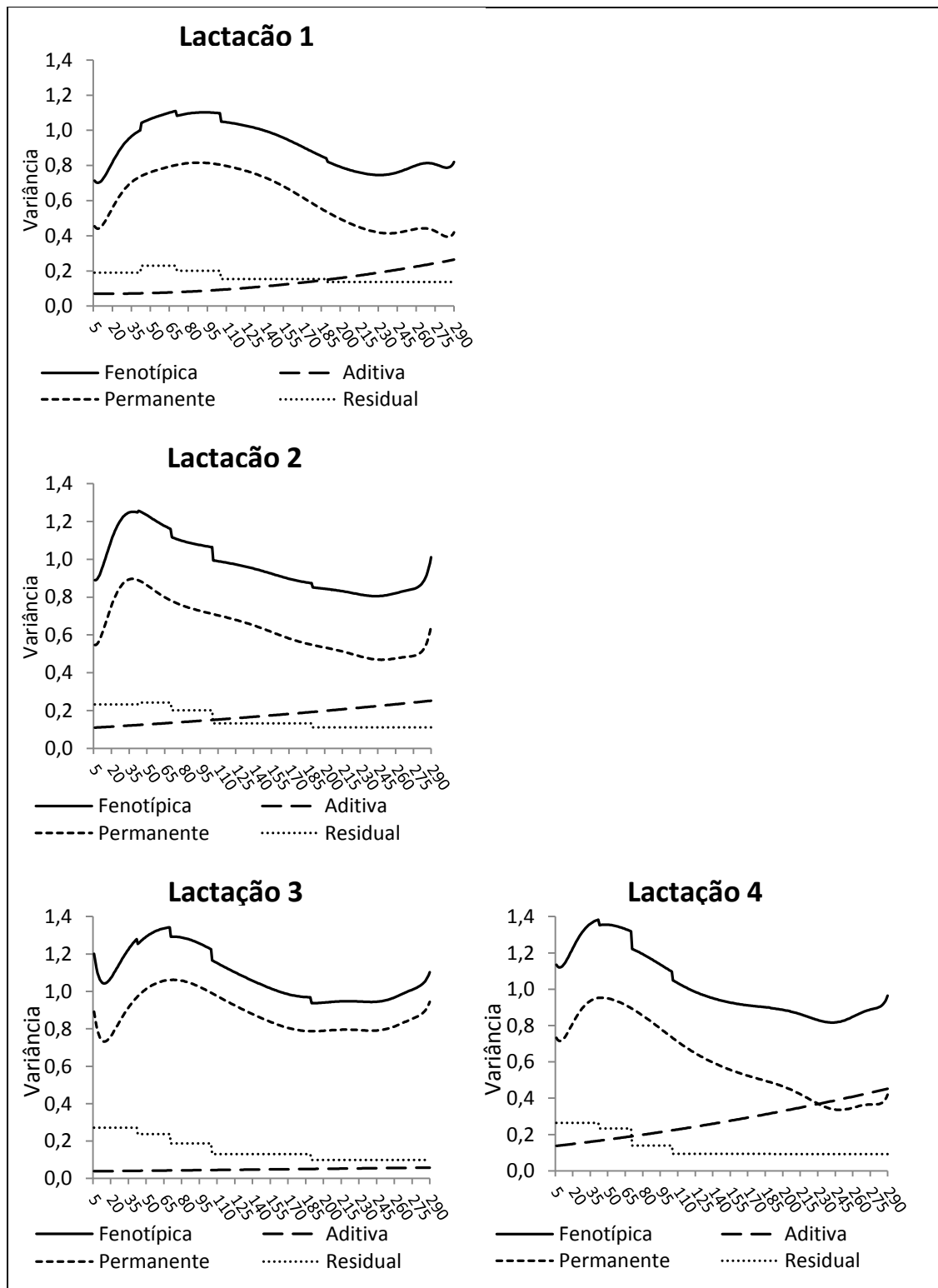


Figura 1: Estimativas das variâncias fenotípica, genética aditiva, permanente individual e residual obtidos em função do número de dias em lactação para as quatro primeiras lactações de caprinos da raça Alpina.

A Figura 2 apresenta as estimativas de herdabilidade ao longo dos dias em lactação para as quatro primeiras lactações, identificadas respectivamente por: L1, L2, L3 e L4. E a porcentagem da variação fenotípica explicada pelas regressões do modelo.

As curvas de herdabilidade das quatro lactações divergiram muito, sendo um pouco mais próximas as curvas da primeira e segunda lactação. Tais divergências podem indicar alterações nas estruturas genéticas relacionada à produção de leite em cada lactação, não necessariamente os genes ligados diretamente a fisiologia da produção, mas principalmente genes de persistência e resistência a doenças, que afetam mais os animais a partir da terceira lactação.

A razão pela qual o pico de lactação muitas vezes possui menores valores de herdabilidade pode estar associada aos efeitos de escala da produção, bem como as alterações muito maiores as quais os animais no pico de lactação sofrem quando atingidos por fatores ambientais desfavoráveis ou favoráveis. Por exemplo, cabras infectadas por doenças, durante o pico da lactação, terão maior queda de produção quando comparadas a animais infectados no final da lactação. Já a melhoria do concentrado fornecido às cabras no início da lactação poderá refletir em um grande aumento na produção dos animais no pico de lactação, mas não fará muita diferença se o animal estiver próximo de encerrar a lactação.

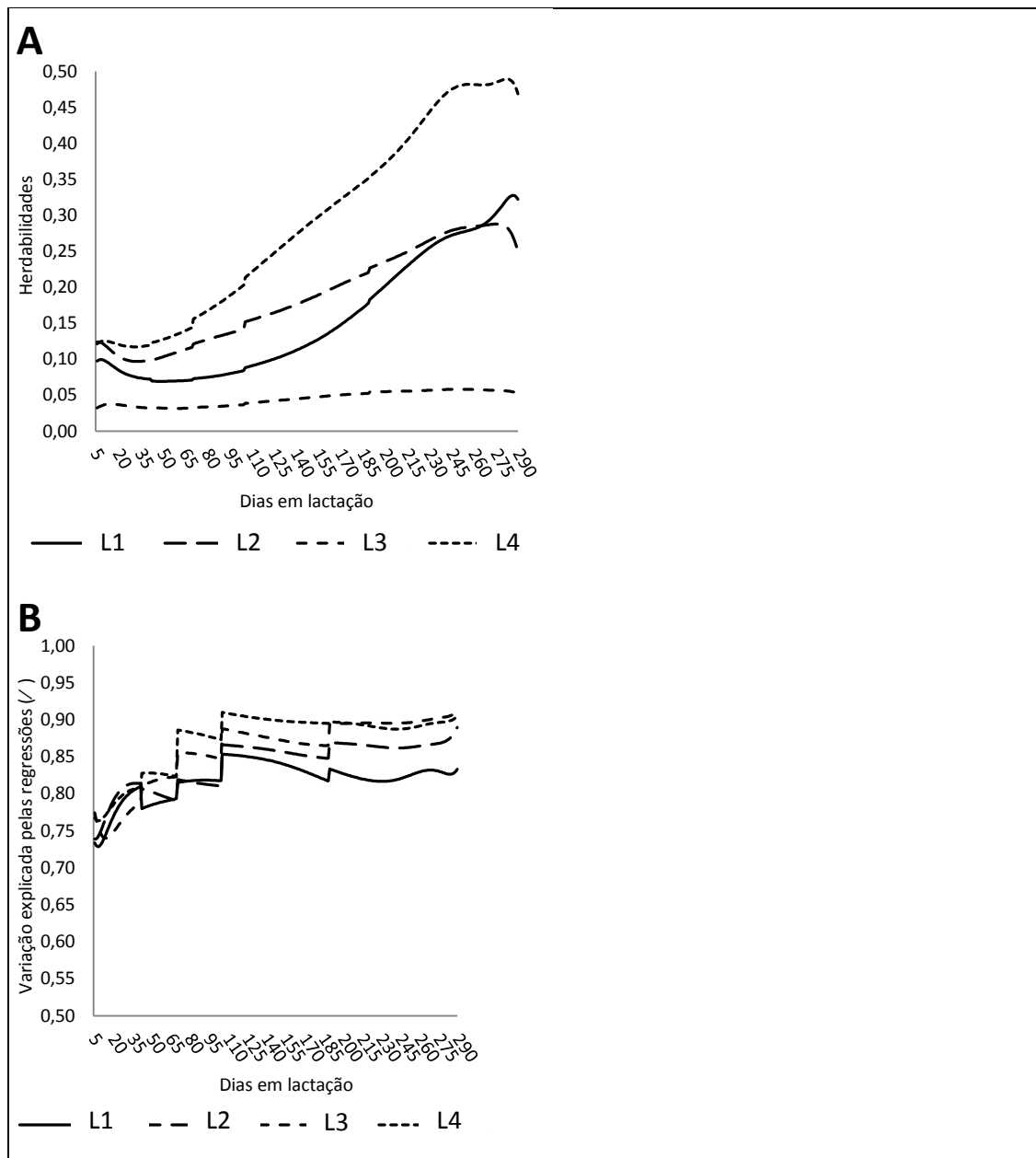


Figura 2: Herdabilidades (A) e porcentagem da variação fenotípica explicadas pelas regressões utilizadas no modelo (B) em função do número de dias em lactação para as quatro primeiras lactações, identificadas de L1 a L4.

As alterações das estimativas também podem estar relacionadas a outros fatores como baixa qualidade das estimativas à medida que aumenta o número da lactação, o que pode ocorrer devido ao número cada vez mais reduzido de observações.

As correlações genéticas entre os diferentes dias em lactação (de 6 a 290) foram todas altas, indicando pouca alteração na estrutura genética que influencia a produção de leite ao longo da lactação, embora a estrutura média da população tenha seu formato característico. Os resultados não foram apresentados graficamente pois a baixa variação tornaria tal representação não informativa, visto que nenhuma correlação, entre dias em lactação, para qualquer das lactações foi inferior a 95%.

Através da Figura 2 é possível perceber que a primeira lactação é a que apresenta menor porcentagem de variação explicada pelas regressões do modelo ao longo dos dias, mesmo sendo o modelo escolhido para a mesma dentre 980 modelos testados por Silva et al. (2013b), isso provavelmente se deve ao fato da primeira lactação termos cabras com diferentes idades ao primeiro parto, e esta covariável com efeitos linear e quadrático não ser o suficiente para corrigir as variações causadas quando a fêmea pari muito cedo, podendo neste caso muitos genes de precocidade produtiva e reprodutiva estarem envolvidos no processo. Já para idade ao segundo, terceiro e quarto parto, espera-se menor divergência da produção de leite no início da lactação em função da idade da cabra ao parto, pois todos animais já desenvolveram o suficiente para passar pelas fases reprodutivas e produtivas.

Também é possível que a dificuldade de modelar os dados no início de todas as lactações seja devido a um efeito de escala em que valores mais altos produzirão maiores dispersões. Além disso, é perceptível que os estresses causados no início da lactação podem provocar alterações maiores não só em quantidade, mas também em proporção, enquanto o final da lactação tem alterações de menor magnitude.

Em rebanhos bovinos, as taxas de reposição utilizadas são de aproximadamente 20%, em caprinos, taxas de reposição maiores são permissíveis devido a maior prolificidade média destes animais, mas isso exigirá também um manejo reprodutivo bem feito, além de um bom trabalho na fase de cria alcançando baixa mortalidade. As taxas de reposição podem aumentar ainda mais se a sazonalidade reprodutiva não for empecilho no rebanho, o que pode ocorrer em latitudes menores do que 20° ou no caso de outras latitudes, utilizando-se programas de luz artificial e aplicação de hormônios para se obter três partos em dois anos.

A Tabela 1 apresenta as estatísticas descritivas para os valores amostrais da PLDC nas quatro lactações. E os valores de LogL, NCA e NCP utilizados como critérios de avaliação do modelo nas diferentes lactações.

Tabela 1: Estatística descritiva para os valores amostrais da PLDC, contendo o número de observações (NO e N%), o menor valor encontrado (min), o valor médio (média), o valor máximo (max) e o desvio padrão (std). E critérios de comparação de modelos nas diferentes lactações (LogL e os número de condição para a regressões genética aditiva e permanente individual, denominados NCA e NCP, respectivamente).

| OP | NO | N% | min | média | max | Std | LogL | NCA | NCP |
|----|--------|-------|-----|-------|-----|--------|----------|--------|--------|
| 1 | 21.610 | 42,89 | 0,1 | 2,20 | 6,9 | 0,9888 | 4351,303 | 13,43 | 182,44 |
| 2 | 14.738 | 29,24 | 0,1 | 2,05 | 7,5 | 1,0736 | 2841,662 | 199,00 | 134,78 |
| 3 | 8.776 | 17,41 | 0,1 | 1,96 | 6,5 | 1,1063 | 1626,664 | Ind. | 362,40 |
| 4 | 5.274 | 10,46 | 0,1 | 1,97 | 7,2 | 1,1002 | 1428,146 | 231,56 | Ind. |

O termo Ind. significa uma indeterminação, que ocorre quando o menor autovalor estimado é muito próximo de zero, pois o menor autovalor encontrado é utilizado no denominador do cálculo do número de condição.

Os valores muito altos ou indeterminados para NCA e NCP podem indicar que o modelo apresenta parâmetros redundantes para a regressão genética aditiva e permanente individual, respectivamente, para a lactação em questão. Conforme apresentado na Tabela 1, isso ocorreu para NCA na terceira lactação e NCP na quarta lactação.

Os valores do LogL caíram muito de uma lactação para outra, o que pode ter sido influenciado pela redução no volume de dados. A redução dos dados ocorreu não apenas no número de animais analisados, mas principalmente da menor proporção de registros por animais, o que compromete diretamente a qualidade de ajustamento de modelos de regressão. Mas não servirá neste caso para contrastar as lactações.

A Tabela 2 apresenta as porcentagens de coincidência, as correlações de Pearson e Spearman e as porcentagens de animais mantidos em comum, após o descarte de 20%, 30% e 40% dos animais de piores valores genéticos.

Tabela 2: Porcentagem de coincidência entre o ordenamento dos valores genéticos (VG), bem como as correlações de Pearson e Spearman para os VG, além das porcentagens de animais mantidos em comum após o descarte de 20%, 30% e 40% dos piores valores genéticos.

| Ordens de Parto | Porcentagem de Coincidência | Correlação de Spearman (%) | Correlação de Pearson (%) | Porcentagem de Indivíduos mantidos em comum quando são descartados: | | |
|-----------------|-----------------------------|----------------------------|---------------------------|---|-------|-------|
| | | | | 20% | 30% | 40% |
| 1 2 | 00,37 | 49,10 | 51,56 | 88,26 | 81,53 | 75,00 |
| 1 3 | 00,51 | 35,59 | 36,23 | 85,22 | 78,68 | 69,43 |
| 1 4 | 00,34 | 41,16 | 41,77 | 84,95 | 77,42 | 72,38 |
| 2 3 | 00,50 | 46,68 | 53,08 | 87,58 | 79,17 | 71,43 |
| 2 4 | 00,25 | 57,93 | 52,76 | 86,33 | 82,08 | 78,66 |
| 3 4 | 01,00 | 31,52 | 30,18 | 84,33 | 76,70 | 68,20 |

Em sistemas com alto grau de tecnologia, maiores descartes podem ser almeçados, o que aumenta a intensidade de seleção podendo gerar aumento do ganho genético. Em rebanhos com menor nível tecnológico, menores níveis de descarte poderão ser realizados.

Os resultados expostos na Tabela 2 demonstram que em menores taxas de descarte a porcentagem de animais mantidos em comum para as quatro lactações estudadas é maior, o que deve ser visto com muito entusiasmo, pois indica que convencer o produtor a realizar o controle leiteiro e avaliação genética ao menos da primeira lactação, poderá permitir a seleção de animais de melhor genética para todas as lactações.

A medida anterior pode ajudar os responsáveis técnicos do melhoramento de rebanhos de pequenos produtores a convencer os mesmos a iniciar a atividade do controle leiteiro ao menos da primeira lactação, podendo obter também ganho nas demais

lactações. Entretanto, em sistemas com maior nível de tecnificação é recomendável investir recursos para se avaliar todas as lactações possíveis obtendo respostas de maior retorno econômico.

Os resultados obtidos por Silva et al. (2013b) de que para sistemas pouco tecnificados poderiam obter boa resposta na seleção se praticassem controles leiteiros com intervalos de no máximo 21 dias também podem ajudar a convencer os produtores, vistos que controles semanais, como é o caso dos realizados no rebanho em estudo, podem ser vistos como grande empecilho por parte dos produtores devido à tendência mundial de aumento do custo de mão de obra qualificada.

De acordo com as correlações de Pearson e Spearman encontradas e principalmente para a porcentagem de coincidência entre a ordem dos indivíduos classificados pelos valores genéticos obtidas da avaliação de cada lactação, não seria adequado recomendar que a confecção de catálogos de reprodutores em hipótese alguma seja feita utilizando apenas informações da primeira lactação. Isto por que a divulgação nos catálogos, dos animais melhor avaliados influencia a utilização dos mesmos. E no caso de se utilizar informações apenas da primeira lactação percebe-se que os melhores reprodutores não serão aqueles que em média produzirão as filhas que promoverão melhor retorno econômico em todas as lactações.

Os resultados mostraram diferenças entre o comportamento das regressões aditivas para cada lactação, sugere-se que também sejam realizados estudos buscando o melhor modelo para cada lactação se o intuito for a realização de avaliações genéticas.

O alto grau de importância e aplicabilidade do assunto deste trabalho nos permite sugerir que futuramente sejam realizados estudos de modelos de regressão aleatória multivariados (MRAM), que trazem consigo vantagens de considerar a estrutura de covariância existente entre as lactações aproveitando melhor os dados e reduzindo problemas de falta de informações em lactações mais avançadas.

Entretanto, muito cuidado deve ser tomado durante estes estudos, visto que MRAM também trazem consigo grandes desvantagens, tais como grande aumento do número de parâmetros a serem estimados pelo modelo e grande dificuldade de convergência. Desta forma estes modelos e rotinas, são muito recentes e pouco testados e podem levar a estimativas parâmetros e previsões de valores genéticos não coerentes com a realidade, o que pode ocasionar equívocos no processo de seleção.

Conclusão

Diferenças entre a estrutura genética de cada uma das lactações devem ser estudadas para delinear adequadamente estratégias de seleção em rebanhos de aptidão leiteira.

O tamanho do rebanho e o nível de recursos financeiros e tecnológicos devem ser combinados com os resultados dos estudos genéticos para prover resultados socioeconômicos mais eficazes.

Futuras análises testando diferentes modelos para a segunda, terceira e quarta lactação e a avaliação da utilização de modelos multicaracterísticos de regressão aleatória envolvendo todas as lactações podem agregar mais informações úteis à discussão dos assuntos tratados neste trabalho.

Referências Bibliográficas

- AKAIKE, H. **Information theory and an extension of the maximum likelihood principle**. Proceedings of 2nd International Symposium on Information Theory, Budapest, Hungary, 1973. p.267-281.
- BRITO, L.F. Modelagem da produção de leite de cabras das raças Alpina e Saanen utilizando Regressão Aleatória. UFV, 2012. 82f. Dissertação (Mestrado), Programa de Pós-Graduação em Genética e Melhoramento (PPGGM), Universidade Federal de Viçosa (UFV). 2012.
- BRITO, L.F.; SILVA, F.G.; MELO, A.L.P.; CAETANO, G.C.; TORRES, R.A.; RODRIGUES, M.T.; MENEZES, G.R.O. Genetic and environmental factors that influence production and quality of milk of Alpine and Saanen goats. **Genetics and Molecular Research**. p.3794-3802, 2010.
- KIRKPATRICK, M.; LOFSVOLD, D.; BULMER, M. Analysis of the inheritance, selection and evolution of growth trajectories. **Genetics**, v.24, n.3, p.979-993, 1990.
- LIDAUER, M.; MANTYSAARI, E.A.; STRADÉN, I.; POSO, J. Multiple-trait random regression test-day model for all lactations. *Interbull bulletin* 25:81. 2000.
- MENEZES, G.R.O.; CAETANO, G.C.; BRITO, L.F.; SILVA, F.G.; MELO, A.L.P.; TORRES, R.A.; RODRIGUES, M.T. Estudo da produção de leite em cabras da raça Saanen utilizando modelos de regressão aleatória. VIII Simpósio Brasileiro de Melhoramento Animal Maringá, PR – 01 e 02 de julho de 2010.

- MEYER, K. (2007). WOMBAT – A tool for mixed model analyses in quantitative genetics by REML. **J. Zhejiang Uni. SCIENCE B** 8: 815–821. [doi:10.1631/jzus.2007.B0815]
- MISZTAL, I. (2008) Reliable computing in estimation of variance components. *Journal of animal breeding and genetics* [0931-2668] vol: 125. Fasc:6 pág: 363-370.
- MONTGOMERY, D. C.; & PECK, E. A. **Introduction to linear regression analysis**. New York, John Wiley & Sons, 1981. 504 p.
- RAO, C.R. **Linear statistical inference and its applications**. 2.ed. New York: John Willy & Sons. p.417-420. 1973.
- SILVA, F. G. **Estudo da produção de leite de caprinos utilizando modelos de regressão aleatória**. Viçosa, MG. 2011. 210 p. Dissertação (Mestrado em Genética e Melhoramento) – Programa de Pós-Graduação em Genética e Melhoramento (PPGGM), Universidade Federal de Viçosa (UFV). 2011.
- SILVA, F. G.; BRITO, L. F.; TORRES, R. A.; RIBEIRO JÚNIOR, J.I.; OLIVEIRA, H. R.; CAETANO, G.C.; RODRIGUES, M. T. Factors that influence the test day milk yield and composition. **Genetics and Molecular Research**. v.12, p.1522 - 1532, 2013a.
- SILVA, F.G.; TORRES, R.A.; BRITO, L.F.; EUCLYDES, R. F.; MELO, A. L. P.; SOUZA, N. O.; RIBEIRO Jr. J.I. and RODRIGUES M.T. Random regression models using Legendre orthogonal polynomials to evaluate the milk production of Alpine goats. **Genetics and Molecular Research**. v.12(4), p.6502-6511, 2013b.

CAPÍTULO 2

Avaliação genética da produção de leite de caprinos da raça Alpina para as três primeiras lactações utilizando modelos de regressão aleatória multicaracterísticos

Resumo: Acredita-se que modelos de regressão aleatória multicaracterísticos (MRAM) apresentem a vantagem de aproveitar melhor as informações de mais de uma lactação simultaneamente, predizendo valores genéticos mais acurados através da estrutura de covariância entre os parâmetros de diferentes lactações. Objetivamos com este trabalho estudar a aplicação de MRAM na avaliação da PLDC das três primeiras lactações de caprinos da raça Alpina. O programa RENPED foi utilizado para realizar a estatística descritiva, a verificação de erros no pedigree, a recodificação dos dados e as correlações de Pearson e Spearman entre os valores genéticos. O programa WOMBAT foi utilizado para estimar componentes de variância e realizar as avaliações genéticas dos modelos de regressão aleatória unicaracterísticos (MRAU) e multicaracterísticos. Os MRAM e MRAU apresentam estimativas semelhantes para os componentes de variância ao longo da curva de lactação. Entretanto a classificação dos melhores animais muda consideravelmente comparando MRAU e MRAM, principalmente para os valores genéticos associados a lactações com menos dados fenotípicos. Acredita-se que a diferença se deva à vantagem de MRAM poderem aproveitar melhor os dados de ao menos duas lactações simultaneamente. Entretanto, a execução dos MRAM é muito mais exigente com relação ao conjunto software e hardware, além de apresentar maior dificuldade para convergir. Mais estudos, inclusive de simulação de dados são necessários para aumentar a relevância desta discussão.

Palavras-chave: curva de lactação, Legendre, valor genético acumulado

Introdução

Segundo Resende (2001), modelos de regressão aleatória (MRA) podem ser considerados como modelos multivariados simplificados ou reduzidos que permitem a obtenção dos mesmos parâmetros de interesse que podem ser obtidos pelos modelos multivariados, mas com uma menor parametrização. De fato, se correta a modelagem, os MRA atenderão adequadamente a este propósito. Silva et al. (2013b) avaliaram 980 MRA utilizando um Polinômio Ortogonal de Legendre (POL), os autores descreveram um polinômio de regressões fixa, genética aditiva e permanente individual de ordens quatro, dois e sete, respectivamente, além de cinco classes da variância residual.

Kistemaker (1997) demonstrou, utilizando estudos de simulação, que os modelos de regressão aleatória avaliando a Produção de Leite no Dia do Controle (PLDC) geravam resultados significativamente melhores que as análises tradicionais da Produção de Leite Acumulada (PLA). Gerando incrementos de acurácia dos valores genéticos dos touros de 2 a 3% e das vacas de 6 a 8%.

Nas avaliações da produção de leite, apesar dos MRA trazerem grandes vantagens através da viabilização da avaliação da PLDC em detrimento da PLA, também existem algumas desvantagens, uma delas é a não possibilidade de aproveitar simultaneamente as informações de mais de uma lactação e a estrutura de covariância entre estas.

Tal problema pode ser resolvido através da utilização de Modelos de Regressão Aleatória Multicaracterísticos (MRAM), dos quais acredita-se que uma das principais vantagens em detrimento de MRA seja a possibilidade de predizer com maior acurácia os valores genéticos para lactações em que alguns animais não tenham informações, o que é frequente por perda de dados e remoção de animais, definitiva ou temporária do plantel.

Jamrozik et al. (1997), testaram a utilização MRAM para analisar a PLDC, a produção de gordura, a produção de proteína e o escore de células somáticas, dentro de três lactações. Os autores chamaram esta tentativa de escolha lógica de teste frente as mudanças do sistema de coletas de dados produtivos do rebanho canadense. Espera-se dos MRAM melhores resultados frente ao aumento de características contínuas de interesse econômico a serem avaliadas em um rebanho.

Objetivou-se com este estudo verificar a aplicação adequada de MRAM para avaliar a PLDC de caprinos para as três primeiras lactações, bem como avaliar as dificuldades de convergência e as complicações para se utilizar os MRAM.

Material e Métodos

Foram utilizados dados da PLDC de cabras do setor de caprinocultura da Universidade Federal de Viçosa (Viçosa, MG). Os animais foram criados em sistema de estabulação livre e receberam alimentação à base de silagem de milho e feno, como volumoso, e mistura concentrada, fornecida conforme a exigência nutricional dos animais.

Os controles leiteiros foram realizados semanalmente em ordenha mecânica, realizada duas vezes ao dia. A coleta na parte da manhã teve início às seis horas e a coleta na parte da tarde teve início às quatorze horas. A PLDC representa a soma do leite coletado nestes dois períodos. Após a análise crítica dos dados restaram 45.131 registros de PLDC de 835 cabras da raça Alpina no arquivo completo.

O software RENPED (Silva, 2011) foi utilizado para corrigir erros de pedigree e recodificar os dados alfanuméricos, procedimentos que melhoram e viabilizam a análise dos dados, respectivamente. O software WOMBAT (Meyer, 2007) foi utilizado em todas as análises genéticas, fornecendo as estimativas dos parâmetros e previsões dos valores genéticos para cada animal avaliado.

O modelo utilizado nas avaliações da PLDC continham combinações de ajuste para os polinômios das curvas, fixa, genética aleatória e de ambiente permanente de ordem 4, 2 e 7, além do número de classes de variâncias residuais igual a 5. Seguindo o melhor modelo encontrado por Silva et al. (2013b), para o mesmo rebanho.

O modelo animal unicaracterístico de regressão aleatória utilizado para avaliar as três primeiras lactações separadamente é descrito de forma generalizada como:

$$y_{ij} = EF_i + \sum_{m=0}^{4-1} b_m \phi_m(t_{ij}) + \sum_{m=0}^{2-1} \alpha_{im} \phi_m(t_{ij}) + \sum_{m=0}^{7-1} \gamma_{im} \phi_m(t_{ij}) + \varepsilon_{ij}$$

, em que: y_{ij} corresponde à produção de leite no controle j da cabra i ; EF_i consiste no conjunto de efeitos fixos, sendo a parte classificatória constituída da média, do grupamento genético (Alpina > 90% = 1; 80% < Alpina ≤ 90% = 2; 70% < Alpina ≤ 80% = 3; 60% < Alpina ≤ 70% = 4; 50% < Alpina ≤ 60% = 5), do ano-estação e tipo de parto. A idade da cabra ao parto foi incluída como covariável, com efeitos linear e quadrático; b_m é o coeficiente de regressão m da PLDC sobre o polinômio de Legendre para modelar a curva média da população; α_{im} e γ_{im} são os m coeficientes de regressão genéticos aditivos e de ambiente permanente, respectivamente, para a cabra i ; 4-1, 2-1 e 7-1 são os graus dos polinômios ortogonais de Legendre; t_{ij} é a variável controle, semana de lactação da cabra i padronizada para o intervalo de -1 a 1; $\phi_m(t_{ij})$ é a função polinomial

de Legendre para o parâmetro m avaliada para a idade t_{ij} ; e ε_{ij} consiste nos efeitos específicos à cada observação não explicados pelos fatores de correção nem pelas regressões do modelo.

Neste trabalho utilizamos os mesmos efeitos fixos encontrados em Silva et al. (2013a), para avaliar a PLDC, por se tratar de dados do mesmo rebanho. Em rebanhos ou características distintas, os autores recomendam a realização de novos testes. Para estudar e entender melhor divergências nos resultados de análises de MRAM em detrimento de MRA será adotado sempre o mesmo, caso contrário poderia haver confundimento entre a estrutura genética e o efeito de modelagem.

O modelo animal multicaracterístico de regressão aleatória utilizado nas avaliações simultâneas de mais de uma lactação é uma extensão dos MRA anteriormente apresentados, as diferenças básicas na estrutura das equações de modelos mistos serão apresentadas a seguir de forma simplificada.

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z & X'R^{-1}W \\ Z'R^{-1}X & Z'R^{-1}Z+A^{-1}\otimes G_0^{-1} & Z'R^{-1}W \\ W'R^{-1}X & W'R^{-1}Z & W'R^{-1}W+I\otimes P^{-1} \end{bmatrix} \cdot \begin{bmatrix} b^0 \\ \hat{a} \\ \hat{p} \\ \sim \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ \sim \\ Z_1'R^{-1}y \\ \sim \\ Z_2'R^{-1}y \\ \sim \end{bmatrix}, \text{ em que:}$$

$$X = \begin{bmatrix} X_1 & \emptyset \\ \emptyset & X_2 \end{bmatrix}, Z = \begin{bmatrix} Z_1 & \emptyset \\ \emptyset & Z_2 \end{bmatrix}, W = \begin{bmatrix} W_1 & \emptyset \\ \emptyset & W_2 \end{bmatrix}, \hat{a} = \begin{bmatrix} \hat{a}_1 \\ \sim \\ \hat{a}_2 \\ \sim \end{bmatrix}, \beta^0 = \begin{bmatrix} \beta_1^0 \\ \sim \\ \beta_2^0 \\ \sim \end{bmatrix}, y = \begin{bmatrix} y_1 \\ \sim \\ y_2 \\ \sim \end{bmatrix}$$

$$G_0 = \begin{bmatrix} G_{01} & \dots & G_{021} \\ \vdots & \ddots & \vdots \\ G_{012} & \dots & G_{02} \end{bmatrix}, Z_1 = \begin{bmatrix} 1 & -1 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & \dots & 0 & 0 & 0 \\ 0 & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & -1 & 1 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 1 & 1 \end{bmatrix} e \hat{a}_1 = \begin{bmatrix} \hat{a}_{110} \\ \hat{a}_{111} \\ \hat{a}_{112} \\ \vdots \\ \hat{a}_{18510} \\ \hat{a}_{18511} \\ \hat{a}_{18512} \end{bmatrix}$$

Crerios de comparaço de modelos tais como logaritmo da funço de verossimilhança restrita (LogL) e crerio da informao de Akaike (AIC, por Akaike, 1973) no podem ser utilizados para comparar os tipos de anlises, principalmente por que o uso de conjuntos de dados completos ou repartidos altera os resultados.

A apresentao do nmero de condio (NC) ser utilizado para verificar problemas de multicolinearidade nos parmetros de estimao do modelo. O teste da rao de verossimilhança no ser utilizado, pois no faz sentido contrastar o MRAM, separadamente, com os quatro MRA com bancos de dados repartidos.

Os resultados dos valores genéticos para os parmetros do modelo aditivo foram integrados, utilizando o software RENPED, para obter o valor genético acumulado de

cada animal em cada lactação e nas quatro lactações utilizando MRAU e MRAM. O software Calc da LibreOffice foi usado para calcular as correlações de Pearson e Spearman entre os valores genéticos acumulados para cada lactação.

Resultados e Discussão

Um dos grandes problemas na execução deste trabalho está relacionado ao alto grau de parametrização dos MRAM, que apresentam dificuldades em convergir, e são muito exigentes quanto ao conjunto software e hardware. Em particular a este estudo destaca-se os altos graus da curva de ambiente permanente definidos por Silva et al. (2013a), os quais inclusive inviabilizaram que mais do que duas lactações fossem analisadas simultaneamente. Dessa forma, sugerimos que esta questão seja revista, não do ponto de vista estatístico, mas funcional.

A Tabela 1 apresenta os resultados das correlações de Pearson e Spearman, acima e abaixo da diagonal principal, respectivamente, entre os valores genéticos preditos acumulados para a primeira lactação. De acordo com os resultados os valores genéticos preditos acumulados para a primeira lactação não variaram muito quando obtidos de uma avaliação da PLDC apenas para a primeira lactação utilizando MRAU (1de1) ou quando foi realizada uma análise da PLDC para duas lactações simultaneamente ou a primeira e a segunda (1de1e2) ou a primeira e a terceira (1de1e3).

Tabela 1: Coorelações de Pearson, acima da diagonal principal e Spearman, abaixo da diagonal principal, para os valores genéticos preditos acumulados da PLDC da primeira lactação.

| | 1de1 | 1de1e2 | 1de1e3 |
|--------|--------|--------|--------|
| 1de1 | - | 94,93/ | 97,47/ |
| 1de1e2 | 94,39/ | - | 93,11/ |
| 1de1e3 | 97,44/ | 92,87/ | - |

Análise da primeira lactação utilizando dados da: Primeira lactação (1de1); primeira e segunda lactações (1de1e2); primeira e terceira lactações (1de1e3).

As Figuras 1, 2 e 3 apresentam as estimativas das variâncias fenotípica, genética aditiva, permanente individual e residual para a PLDC da primeira lactação de caprinos da raça Alpina. A Figura 1 apresenta estes resultados com base na análise de um MRAU avaliando apenas a primeira lactação, enquanto a Figura 2 apresenta resultados provenientes da avaliação simultânea das duas primeiras lactações utilizando MRAM, e por fim a Figura 3 o faz considerando a avaliação simultânea das lactações 1 e 3 utilizando também um MRAM.

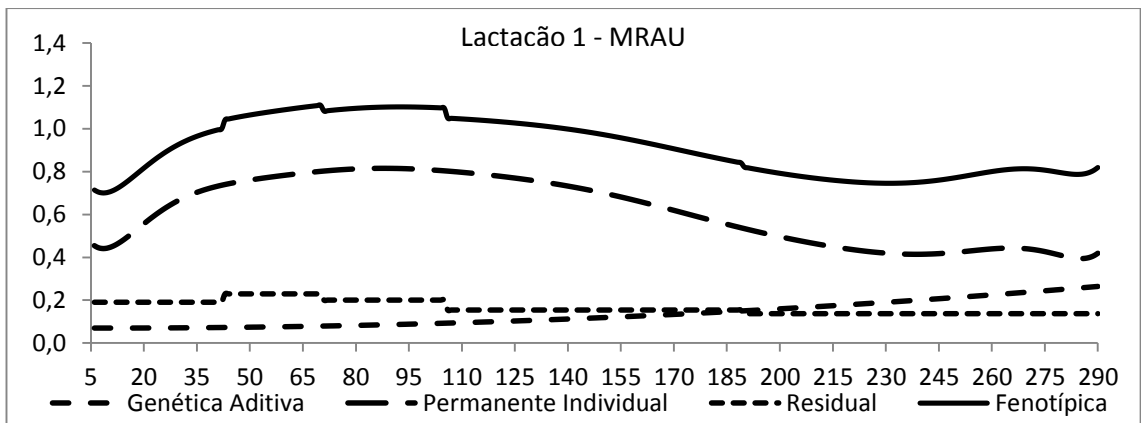


Figure 1: Variâncias fenotípica, genética aditiva, de ambiente permanente e residual para a PLDC para a primeira lactação. Resultados obtidos a partir da avaliação apenas da primeira lactação utilizando MRAU.

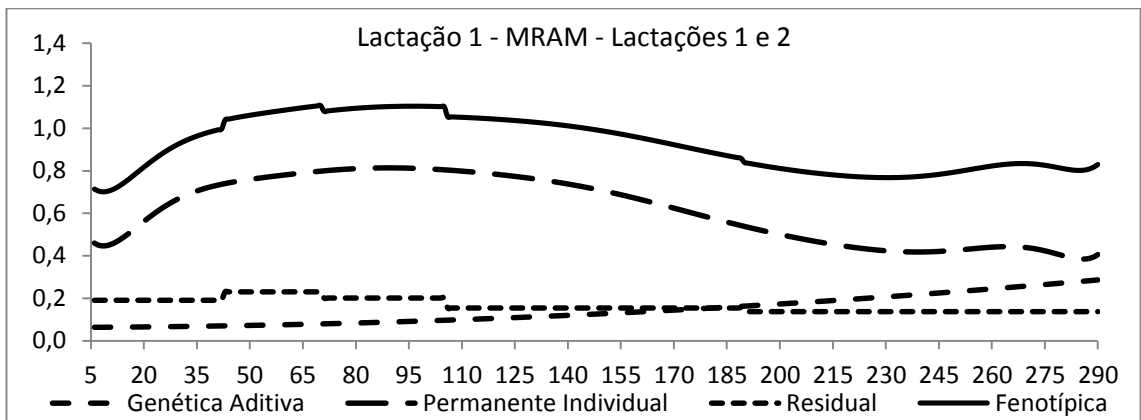


Figure 2: Variâncias fenotípica, genética aditiva, de ambiente permanente e residual para a PLDC para a primeira lactação. Resultados obtidos a partir da avaliação simultânea da primeira e segunda lactações utilizando MRAM.

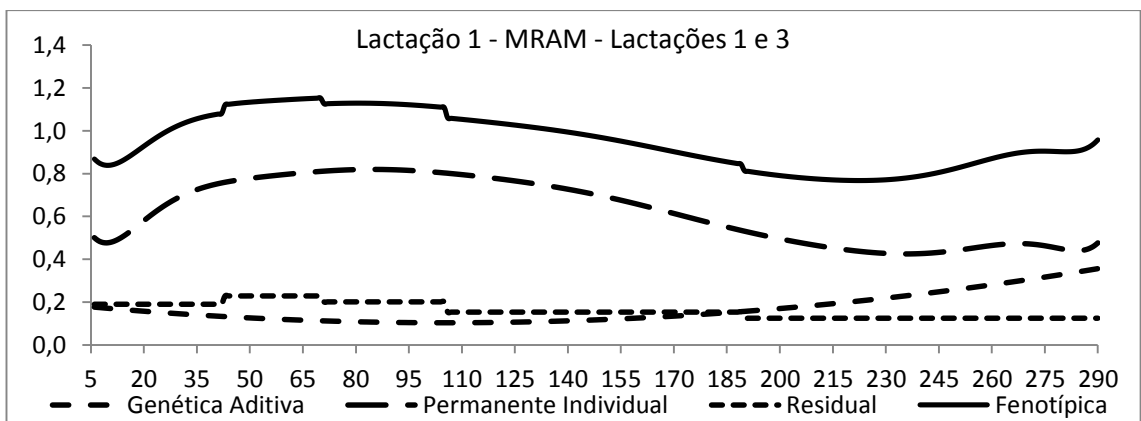


Figure 3: Variâncias fenotípica, genética aditiva, de ambiente permanente e residual para a PLDC para a primeira lactação. Resultados obtidos a partir da avaliação simultânea da primeira e terceira lactações utilizando MRAM.

A Tabela 2 apresenta os resultados das correlações de Pearson e Spearman, acima e abaixo da diagonal principal, respectivamente, entre os valores genéticos preditos acumulados para a segunda lactação. De acordo com os resultados os valores genéticos preditos acumulados para a segunda lactação variaram muito quando obtidos de uma avaliação da PLDC apenas para a segunda lactação utilizando MRAU ou quando foi

realizada uma análise da PLDC para duas lactações simultaneamente, a primeira e a segunda, variando menos quando comparado a segunda e terceira.

Tabela 2: Coorelações de Pearson, acima da diagonal principal e Spearman, abaixo da diagonal principal, para os valores genéticos preditos acumulados da PLDC da segunda lactação.

| | 2de2 | 2de1e2 | 2de2e3 |
|--------|--------|--------|--------|
| 2de2 | - | 20,60/ | 72,57/ |
| 2de1e2 | 88,12/ | - | 30,02/ |
| 2de2e3 | 96,76/ | 99,65/ | - |

Análise da segunda lactação utilizando dados da: Segunda lactação (2de2); primeira e segunda lactações (2de1e2); segunda e terceira lactações (2de2e3).

As Figuras 4, 5 e 6 apresentam as estimativas das variâncias fenotípica, genética aditiva, permanente individual e residual para a PLDC da segunda lactação de caprinos da raça Alpina. A Figura 4 apresenta estes resultados com base na análise de um MRAU avaliando apenas a segunda lactação, enquanto a Figura 5 apresenta resultados provenientes da avaliação simultânea das duas primeiras lactações utilizando MRAM, e por fim a Figura 6 o faz considerando a avaliação simultânea das lactações 2 e 3 utilizando também um MRAM.

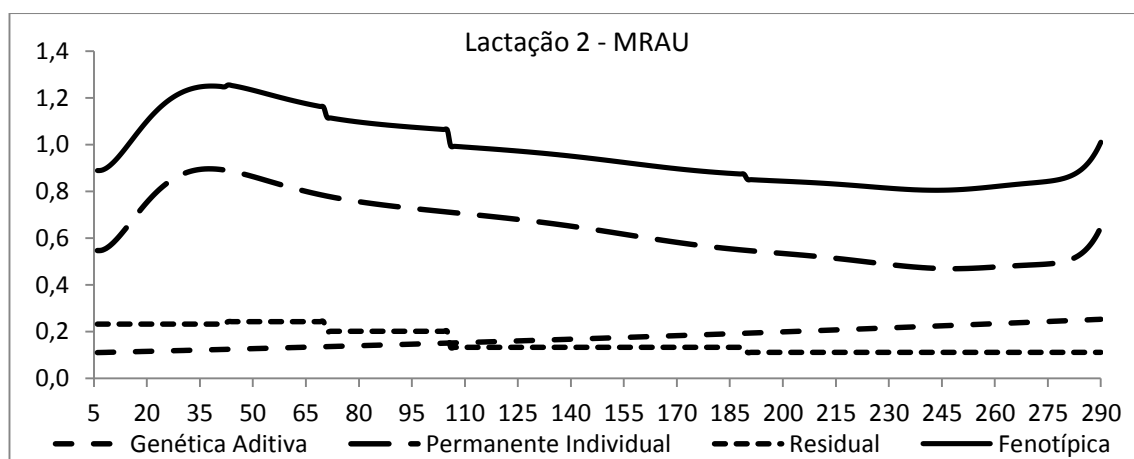


Figure 4: Variâncias fenotípica, genética aditiva, de ambiente permanente e residual para a PLDC para a segunda lactação. Resultados obtidos a partir da avaliação apenas da segunda lactação utilizando MRAU.

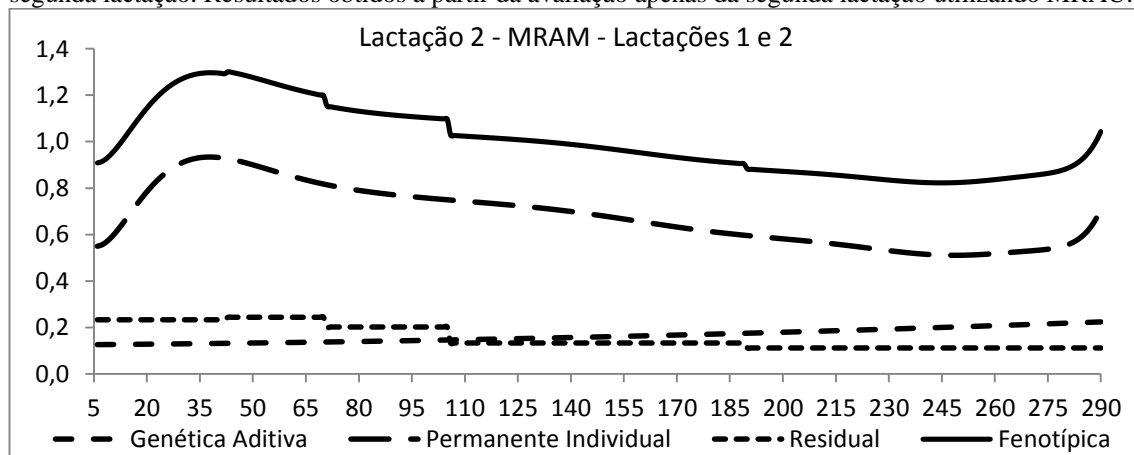


Figure 5: Variâncias fenotípica, genética aditiva, de ambiente permanente e residual para a PLDC para a segunda lactação. Resultados obtidos a partir da avaliação simultânea da primeira e segunda lactações utilizando MRAM.

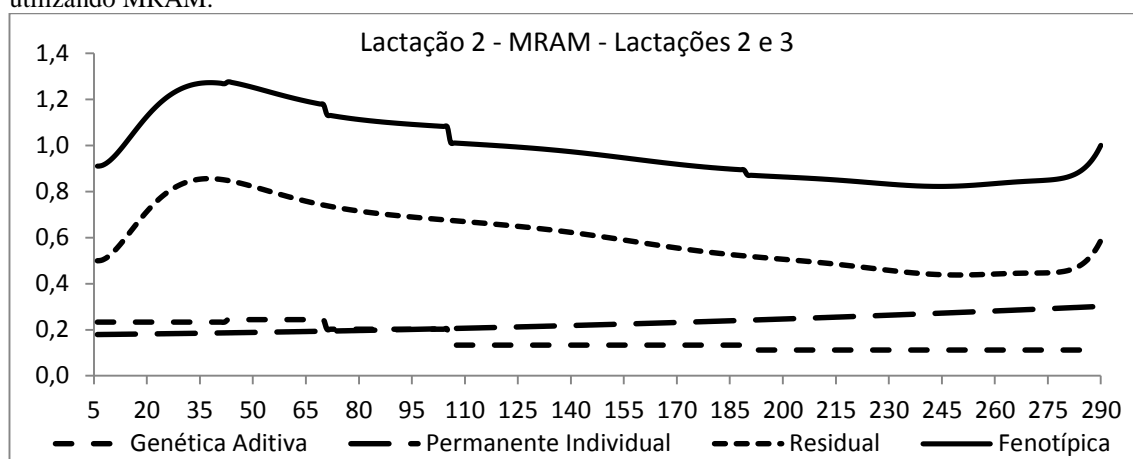


Figure 6: Variâncias fenotípica, genética aditiva, de ambiente permanente e residual para a PLDC para a segunda lactação. Resultados obtidos a partir da avaliação simultânea da segunda e terceira lactações utilizando MRAM.

A Tabela 3 apresenta os resultados das correlações de Pearson e Spearman, acima e abaixo da diagonal principal, respectivamente, entre os valores genéticos preditos acumulados para a terceira lactação. De acordo com os resultados os valores genéticos preditos acumulados para a terceira lactação variam muito quando obtidos de uma avaliação da PLDC apenas para a terceira lactação utilizando MRAU em comparação à avaliações simultâneas para duas, por exemplo a segunda e a terceira, variando menos quando comparado à primeira e terceira.

Tabela 3: Coorelações de Pearson, acima da diagonal principal e Spearman, abaixo da diagonal principal, para os valores genéticos preditos acumulados da PLDC da terceira lactação.

| | 3de3 | 3de1e3 | 3de2e3 |
|--------|--------|--------|--------|
| 3de3 | - | 70,55/ | 12,88/ |
| 3de1e3 | 64,62/ | - | 10,38/ |
| 3de2e3 | 62,86/ | 99,65/ | - |

Análise da terceira lactação utilizando dados da: Terceira lactação (3de3); primeira e terceira lactações (3de1e3); segunda e terceira lactações (3de2e3).

As Figuras 7, 8 e 9 apresentam as estimativas das variâncias fenotípica, genética aditiva, permanente individual e residual para a PLDC da terceira lactação de caprinos da raça Alpina. A Figura 7 apresenta estes resultados com base na análise de um MRAU avaliando apenas a terceira lactação, enquanto a Figura 8 apresenta resultados provenientes da avaliação simultânea das lactações 1 e 3, utilizando MRAM, e por fim a Figura 9 o faz considerando a avaliação simultânea das lactações 2 e 3, utilizando também um MRAM.

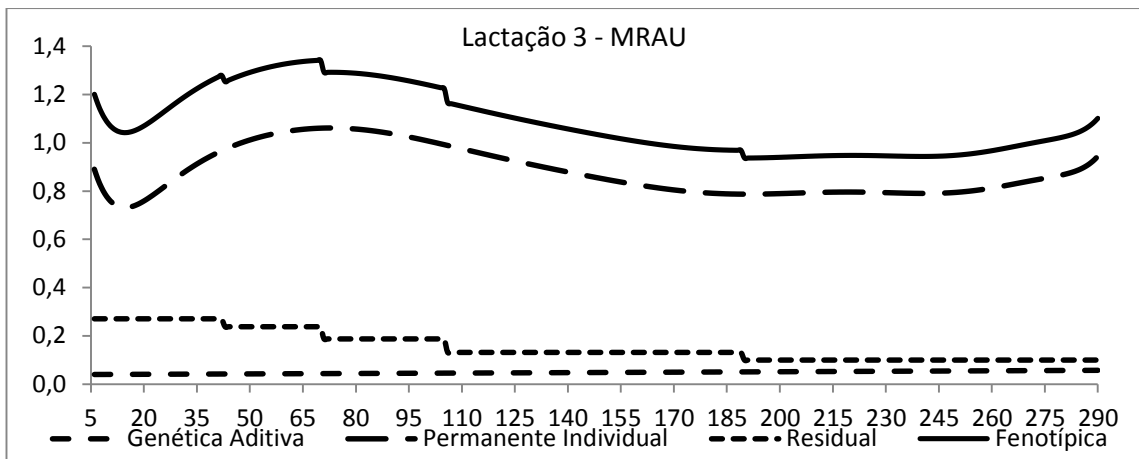


Figure 7: Variâncias fenotípica, genética aditiva, de ambiente permanente e residual para a PLDC para a terceira lactação. Resultados obtidos a partir da avaliação apenas da terceira lactação utilizando MRAU.

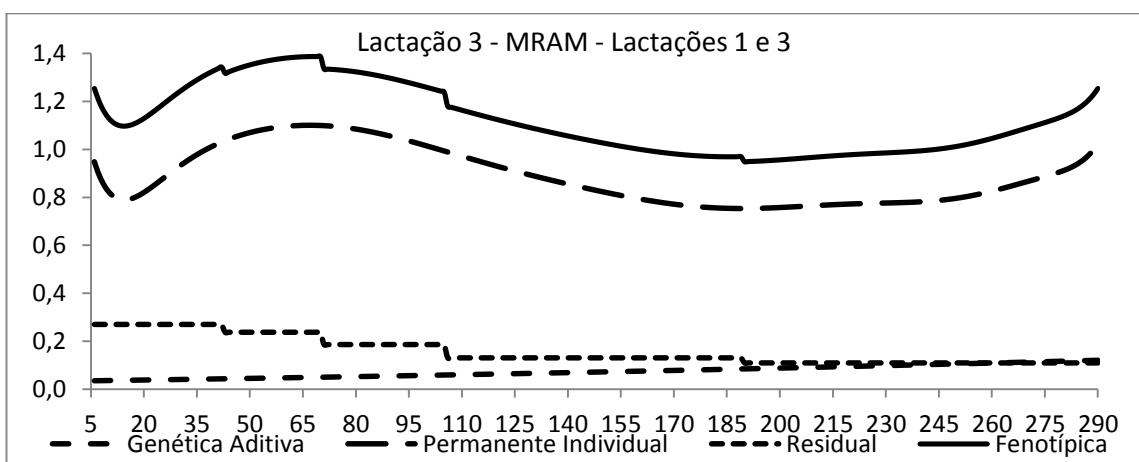


Figure 8: Variâncias fenotípica, genética aditiva, de ambiente permanente e residual para a PLDC para a terceira lactação. Resultados obtidos a partir da avaliação simultânea da primeira e terceira lactações utilizando MRAM.

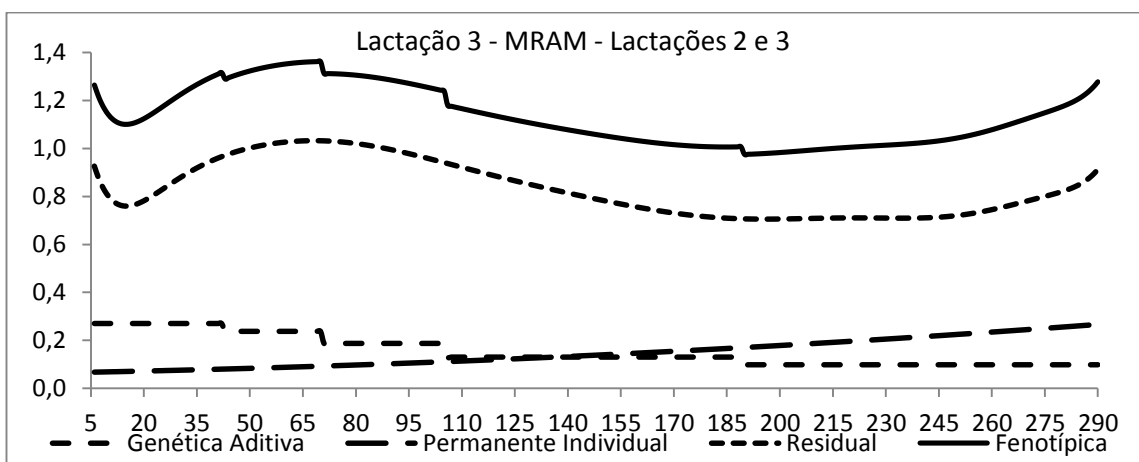


Figure 9: Variâncias fenotípica, genética aditiva, de ambiente permanente e residual para a PLDC para a terceira lactação. Resultados obtidos a partir da avaliação simultânea da segunda e terceira lactações utilizando MRAM.

Os resultados das Figuras 1 até 9 mostraram que as variâncias fenotípicas, genética aditivas, permanente individuais e residuais não se alteraram muito conforme a

utilização de MRAM em detrimento de MRAU. As maiores diferenças foram encontradas como uma distinção da estrutura de variância entre diferentes lactações.

De certo modo estes resultados podem ser vistos de forma positiva, pois não encontrou-se nenhum efeito que pudesse levantar a questão de MRAU ou MRAM poderem estar sendo aplicados de forma errada por levarem a uma modelagem diferente. Os MRAM apresentam maior dificuldade para alcançar convergência e para encontrar valores de partida razoáveis para as análises.

Por outro lado, aliada à informação de algumas correlações de Pearson e Spearman baixas nas Tabelas 2 e 3, estes resultados podem indicar que o baixo volume de dados nas lactações 2 e 3 podem estar influenciando predições de valores genéticos de baixa qualidade, tornando vantajosa a utilização de mais estudos visando a implementação de MRAM nas avaliações genéticas de caprinos de aptidão leiteira.

Uma vez ampliada a qualidade das avaliações de mais de uma lactação, através da avaliação simultânea utilizando MRAM, acredita-se que isso possa permitir a seleção mais eficiente de indivíduos mais equilibrados e produtivos ao longo de toda vida.

Conclusões

MRAM e MRAU apresentam estimativas semelhantes para os componentes de variância ao longo da curva de lactação. Mas ocorreram grandes mudanças na classificações dos animais com relação às análises de cada modelo. O MRAM também apresentaram maior dificuldade de convergência e de obtenção de bons valores de partida para as análises.

Acredita-se que a diferença maior entre os valores genéticos preditos utilizando MRAM e MRAU em lactações com menor volume de dados deve-se à vantagem dos MRAM aproveitarem melhor os dados de ao menos duas lactações simultaneamente. Sugere-se mais estudos, inclusive de simulação de dados, para aumentar a relevância desta discussão.

Referências Bibliográficas

AKAIKE, H. **Information theory and an extension of the maximum likelihood principle**. Proceedings of 2nd International Symposium on Information Theory, Budapest, Hungary, 1973. p.267-281.

- JAMROZIK, J.; SCHAEFFER, L. R.; LIU, Z. & JANSEN, G. 1997. Multiple trait random regression test day model for production traits. Vienna, Austria, Aug. 28-29, 1997. Interbull Bulletin, No. 16, 43.
- KISTEMAKER, G. J. 1997. PhD Thesis. University of Guelph.
- MEYER, K. (2007). WOMBAT – A tool for mixed model analyses in quantitative genetics by REML. J. Zhejiang Uni. SCIENCE B. 8:815–821.
- RESENDE, M. D. V.; RESENDE, G. D. S. P.; FERNANDES, J. S. C. Regressão aleatória e funções de covariância na análise de medidas repetidas. **Revista Mat.** São Paulo, 19, p.21-40, 2001.
- SILVA, F. G. **Estudo da curva de produção de leite de caprinos utilizando modelos de regressão aleatória.** 2011, 144p. Dissertação (Mestrado em Genética e Melhoramento) – Universidade Federal de Viçosa, Viçosa – MG. 2011.
- SILVA, F. G.; BRITO, L. F.; TORRES, R. A.; RIBEIRO JÚNIOR, J.I.; OLIVEIRA, H. R.; CAETANO, G.C.; RODRIGUES, M. T. (2013a) Factors that influence the test day milk yield and composition. **Genetics and Molecular Research.** 12:1522-1532.
- SILVA, F.G.; TORRES, R.A.; BRITO, L.F.; EUCLYDES, R. F.; MELO, A. L. P.; SOUZA, N. O.; RIBEIRO Jr. J.I. and RODRIGUES M.T. (2013b) Random regression models using Legendre orthogonal polynomials to evaluate the milk production of Alpine goats. **Genetics and Molecular Research.** 12:6502-6511.

CAPÍTULO 3

RENPED - Uma ferramenta de auxílio a avaliações genéticas e estimação de componentes de variância

Resumo: A maior parte dos softwares utilizados para avaliações genética de animais ou para a estimação de parâmetros genéticos utilizando a metodologia de modelos mistos de Henderson exige que o arquivo de pedigree contenha apenas dados não alfa numérico, que os animais estejam ordenados e que o número de identificação do animal seja sempre maior que o dos pais. Objetivou-se com este trabalho apresentar o software RENPED e suas utilidades à comunidade científica da área de genética e melhoramento animal através de exemplos aplicados. Foram utilizados 50.398 registros de produção de leite no dia do controle (PLDC) de 2.946 lactações de 1.093 cabras da raça Alpina, e um arquivo de pedigree contendo 7.800 animais. Primeiramente, o arquivo de pedigree foi submetido à verificação de erros e recodificação das informações individuais. Em seguida, o banco de dados foi submetido à recodificação com base nos códigos criados na recodificação do pedigree. O programa WOMBAT foi utilizado para estimar componentes de variância para PLDC e valores genéticos para os animais utilizando modelos de regressão aleatória. As predições encontradas para os parâmetros das curvas genética aditivas dos animais avaliados foram submetidas ao RENPED para integrar a área abaixo da curva genética de cada animal predizendo o valor genético acumulado para a lactação. Os resultados foram posteriormente inseridos no programa RENPED para descodificação e associação dos valores genéticos encontrados às identificações originais dos animais.

Palavras-chave: C++, correspondência, programação, renumeração, relatório de erros

Introdução

Para avaliações da endogamia e programas de direcionamento de acasalamento é necessária a obtenção do parentesco entre os animais. Para tal, geralmente são utilizados algoritmos recursivos que permitam obter a matriz dos numeradores dos coeficientes de parentesco de Wright (1922) (**NRM**). O primeiro método recursivo criado para esse propósito foi desenvolvido por Emik e Terril (1949), denominado método tabular.

A avaliação genética de animais ou estimação de parâmetros genéticos na maior parte das vezes é realizada utilizando-se a metodologia de modelos mistos de Henderson adaptada a diversos modelos por meio de algoritmos computacionais cada vez mais eficientes. Uma característica presente em tais algoritmos é a obtenção direta da inversa da NRM. Henderson (1976) desenvolveu um algoritmo que funcionava de forma eficiente em pedigrees sem indivíduos endogâmicos. Quaas (1976) desenvolveu um artifício para que o algoritmo desenvolvido por Henderson fosse mais eficiente no caso de pedigrees com animais endogâmicos.

Ambas as atividades denotadas anteriormente exigem prévia organização do arquivo de pedigree, demandando não só o ordenamento crescente dos dados pela variável identificação do animal, como também exigem que a numeração dos filhos seja sempre superior à dos pais.

Alguns softwares, como o MTDFREML (Boldman et al., 1993), o DFREML (Meyer, 1998) e o WOMBAT (Meyer, 2007), conseguem realizar estas recodificações e reordenações previamente às etapas de cálculo da matriz de parentesco ou sua inversa. Entretanto estes softwares não trabalham com dados alfa numéricos, que são dados que possuem não só números como também letras ou símbolos, situação frequente ao se trabalhar com animais de médio e grande porte, que muitas vezes recebem nomes ao invés de números.

Outra questão comum em softwares que realizem renumeração é corrigirem a maior parte dos erros automaticamente contendo rotinas que dão preferência a um determinado tipo de correção, o que pode levar a uma escolha inadequada, ou então podem excluir informações indesejadas, o que pode reduzir as informações de parentesco, prejudicando ligações entre famílias. Desta forma, é interessante que para alguns tipos de erros a correção não seja automática, e sim que o software indique o local e o tipo de erros para que o usuário decida como corrigi-los, de forma a acelerar o processo de correção. Neste caso o usuário poderia optar por excluir uma linha de dados, corrigir uma informação

digitada incorretamente, dentre alternativas que exigem o senso crítico do pesquisador e seu conhecimento do banco de dados e suas peculiaridades.

A maioria dos softwares que recodificam pedigrees tem outro problema que é trabalhar apenas com o arquivo de pedigree, exigindo que o pesquisador utilize seus arquivos de saída para recodificar o arquivo de dados utilizando programas de edição de dados, o que aumenta a dificuldade de uso e o tempo de trabalho dos usuários.

Outro uso de pedigrees recodificados é para análises de endogamia. O programa ENDOG (Gutiérrez & Goyache, 2005), por exemplo, exige que os dados estejam recodificados e reordenados, ou ao menos recodificados e com a informação de data de nascimento para todos os animais, o que nem sempre é possível.

Uma alternativa para esses problemas é a utilização do software RENPED, que recodifica e renumera dados do pedigree, aceita trabalhar com dados alfa numéricos, corrige apenas o que não necessita do aval do usuário, realiza a localização de erros que devem ser corrigidos pelo usuário, facilitando a correção. O RENPED também é prático e fácil de utilizar e está disponível em quatro idiomas: Inglês, português, espanhol e francês.

Outra vantagem das recodificações que não pode ser desconsiderada é a vantagem de viabilizarem avaliações genéticas sigilosas, e evitarem direcionamento de sêmen em testes de progênie, e o RENPED também pode ser usado com o propósito de gerar esta recodificação e decodificar quando necessário.

Objetivou-se com este trabalho apresentar o software RENPED e suas utilidades à comunidade científica da área de genética e melhoramento animal através de exemplos aplicados.

Caracterização

Interface do usuário

A interface do programa RENPED esta em ambiente DOS, podendo ser acessado de dentro do sistema operacional Windows. O programa fornece sequências de menus e perguntas para o usuário. Os questionários são simples e as informações requisitadas são fáceis de responder.

O programa gera arquivos de resposta na mesma pasta em que o executável foi acionado, sempre em formato '.txt', desta forma os arquivos de correspondência, erros, dentre outros, podem ser lidos utilizando o bloco de notas.

Aplicações

O RENPED pode ser utilizado para verificar erros e recodificar pedigrees, inclusive contendo dados alfa-numéricos. O programa também permite recodificar os dados para viabilizar avaliações genéticas ou de endogamia, sendo vantagem o seu uso por ser uma ferramenta que permite retornar os dados de resultados de avaliações genéticas ou de endogamia para os dados originais, facilitando a prática de seleção.

O RENPED soluciona ainda um dos maiores problemas da utilização de modelos de regressão aleatória na avaliação de dados longitudinais discretos, que consiste em como selecionar os animais com base nos valores encontrados para as curvas genéticas aditivas. O RENPED calcular rapidamente a soma de todos os valores genéticos preditos para cada dia do controle por animal, e usar este valor como o critério de seleção.

Arquivos de Entrada

O usuário pode escolher o nome e extensão dos arquivos de pedigree, dados e respostas. Entretanto, todos os arquivos devem estar na mesma pasta do executável do programa. Os arquivos devem conter cabeçalho e serem formatados com dados colunados, sendo as colunas de dados separadas por espaço. O RENPED aceita arquivos de entrada nos formatos '.txt', '.prn' ou '.dat'. O arquivo CORRESP.txt é um arquivo de saída e entrada, pois o programa gera o mesmo na primeira opção de execução e o reutiliza, obrigatoriamente, na segunda e terceira opções.

Nos arquivos de entrada, dados perdidos não podem ser deixados em branco, pois neste caso seriam lidos na coluna seguinte, o que acarretaria problema na sequência de leitura. O RENPED permite que dados perdidos sejam representados por números, caracteres ou sequências de caracteres sem espaço. Entretanto uma única escolha deve ser tomada para todo o arquivo e deve ser informada ao programa pelo usuário quando lhe for feita a pergunta: Qual o valor para dados perdidos? Nos arquivos de saída o RENPED padroniza zero para informações de animal, pai ou mãe perdidas.

O arquivo de pedigree pode conter, além das informações de animal, pai e mãe, as informações de sexo e data de nascimento, o que contribui para a verificação de erros. Outras colunas de dados podem ser adicionadas ao pedigree, inclusive as informações do pedigree e banco de dados podem ser incluídas em um único arquivo. Embora isso reduza a eficiência do programa, podendo inclusive limitar o seu uso por insuficiência de memória RAM disponível no computador.

O arquivo de respostas pode conter ou não às informações de pai e mãe, mas obrigatoriamente deve conter a informação de animal. Caso contenha, é importante que

estas sejam declaradas para que as informações dos pais também sejam recodificadas. Espera-se que tal arquivo seja fornecido em formato de texto, proveniente da execução de algum software de avaliação genética o qual se deseja decodificar os animais para atribuir o valor genético ao indivíduo correto na população.

Opções de Configuração

O RENPED possui duas opções de configuração (SETUP OPTIONS), sendo que a primeira opção permite que o usuário modifique o idioma, existindo quatro alternativas: Inglês, português, espanhol e francês. A segunda opção fornece oito tipos de combinações de cor de fundo e texto diferentes do padrão (preto/branco), conforme informado na Tabela 1.

Tabela 1: Relação das opções de combinação de cor de fundo/texto no programa RENPED.

| Opção | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Padrão |
|-------|--------|--------|--------|----------|-------|------------|-------|-------|--------|
| Fundo | Branco | Branco | Branco | Preto | Roxo | Verde água | Preto | Preto | Preto |
| Texto | Azul | Preto | Verde | Vermelho | Verde | Amarelo | Azul | Rosa | Branco |

Opções de Execução

Discutiremos neste trabalho quatro opções principais de execução, a primeira (PEDIGREE RECODER) recodifica pedigrees, a segunda (DATA RECODER) recodifica os dados utilizando o arquivo de correspondências gerado após a execução da primeira opção, a terceira (RESULT DECODER) decodifica arquivos de resultados de análises de endogamia ou avaliação genética de forma a se atribuir o resultado ao código original do animal e também utiliza o arquivo de correspondências gerado após a execução da primeira opção. A quarta opção (RR APPLICATIONS), calcula a área abaixo da curva genética aditiva predita para cada animal, permitindo obter um valor genético acumulado para todo o período de produção.

Ao escolher a opção um será exigido do usuário a informação do nome e extensão do arquivo de pedigree, além da posição das colunas de animal, pai, mãe, sexo e data de nascimento. Entretanto, apenas a coluna com a identificação do animal é obrigatória. Na ausência de qualquer uma das demais colunas o usuário pode utilizar o valor zero no lugar do número da coluna para indicar ao programa que tal informação não está presente no banco de dados.

O programa exige também que o usuário informe o número total de colunas no arquivo de pedigree, não apenas para otimizar o programa reduzindo o espaço alocado na memória durante a execução, assim como para organizar a sequência de leitura do arquivo

de dados. A última informação requisitada pelo programa é o número, caractere ou sequência de caracteres, sem espaço, que identifica dados perdidos.

Ao escolher a opção dois será exigido do usuário a informação do nome e extensão do arquivo de dados, além da posição das colunas de animal, pai e mãe. Entretanto, apenas a coluna com a informação de animal é obrigatória e na ausência das colunas de pais o usuário pode fornecer o valor zero no lugar do número da coluna. Nesta opção também é necessário que o usuário informe o número total de colunas no arquivo de dados.

Ao escolher a opção três será exigido do usuário a informação do nome e extensão do arquivo de resultados, além da posição das colunas de animal, pai e mãe. Entretanto, apenas a coluna com a informação de animal é obrigatória. Na ausência das colunas de pais o usuário deve fornecer o valor zero. Além disso, é necessário que o usuário informe o número total de colunas no arquivo de dados.

Após fornecer as informações o usuário poderá acompanhar o andamento da execução através da barra de loading. O programa também emite um relatório informando se a execução ocorreu com sucesso e informa o número de animais e o número de linhas e colunas no arquivo de dados. Além disso, ele mostra o nome dos arquivos de saída gerados e a utilidade de cada um.

Atualmente modelos de regressão aleatória estão sendo amplamente utilizados para avaliação genética de dados longitudinais, tais modelos fornecem as herdabilidades em cada ponto da curva e permitem aproveitar melhor as informações de produção, não exigindo ajuste ou estimação dos valores previamente às análises.

Para curvas de crescimento, permite a seleção de indivíduos em qualquer ponto dentro do intervalo analisado. Entretanto, para características não cumulativas como PLDC, a seleção em um único ponto da curva de lactação pode acarretar em problemas de perda de resposta correlacionada dos demais períodos ao longo das gerações.

Existe uma alternativa para contornar tal problema, que consiste em se obter os valores genéticos de todas PLDC dentro do intervalo avaliado e somar os resultados, assim passa-se a ter um valor em unidades de medida que equivale à vantagem genética do animal em relação à população analisada, não mais quanto a um dia específico, mas em relação ao período analisado que pode compreender toda a lactação, sendo que metade deste valor corresponderá à diferença esperada na progênie (DEP). Isso pode ser calculado através da opção quatro do RENPED.

O usuário precisa fornecer um arquivo com cabeçalho, contendo as informações de identificação do animal, parâmetro da curva (1, 2, ..., ordem) e valor genético para o parâmetro da curva ou pode simplesmente colocar o arquivo de saída do WOMBAT

(RnSoln_animal.dat) e informar ao RENPED a sub opção número dois. Posteriormente o programa pedirá ao usuário que informe os valores máximo e mínimo para a variável controle (por exemplo: 305 e 5 para lactação de bovinos, considerando a PLDC ou 52 e 6 considerando a produção de ovos semanais). Por fim o programa alerta o usuário para abrir o arquivo INTEGRATED.dat para visualizar os resultados.

Arquivos de Saída

No RENPED a opção de execução um gera três arquivos de saída, sendo estes o arquivo de relatório de erros (REPORT), o arquivo de correspondências entre o pedigree original e o recodificado (CORRESP) e o arquivo de pedigree recodificado (PED-RECOD). A opção dois gera apenas um arquivo de saída, o arquivo de dados recodificado (DATA-RECOD). A opção três também gera apenas um arquivo de saída e o arquivo de resultados decodificado (RES-DECOD). Todos os arquivos de saída possuem extensão ‘.txt’ e são gerados na mesma pasta do executável do programa.

O relatório de erros não só indica os tipos de erros como também informa a localização dos mesmos para que estes possam ser corrigidos pelo usuário, embora alguns sejam corrigidos automaticamente pelo programa.

O primeiro erro relatado consiste em animais que aparecem tanto na coluna de pai quanto na coluna de mãe. O programa informará o ‘nome’ do indivíduo, a linha que este aparece como pai e a linha que aparece como mãe, cabendo ao usuário conferir o tipo de erro ocorrido.

O segundo tipo consiste em machos que aparecem na coluna de mãe, neste caso o relatório indicará a linha onde o animal foi declarado como macho e a linha onde o animal aparece como mãe. O terceiro tipo é o contrário do segundo, e funciona de forma análoga.

O quarto tipo de erro é quando um animal aparecer como pai ou mãe dele mesmo, o programa indicará o ‘nome’ do animal e linha em que o erro se encontra.

O quinto tipo de erro consiste de animais sem informações de sexo, o que pode ser fundamental quando se utiliza tal característica como efeito fixo em uma análise ou deseja-se analisar os dados no programa ENDOG, que exige tal informação para todos os animais. Caso o usuário não possua a informação de sexo no pedigree, ou por algum motivo, não queira utilizá-la, tais erros simplesmente não serão corrigidos.

O sexto tipo de erro consiste de animais sem informação de data de nascimento, o que pode ser fundamental no cálculo de covariáveis como idade ao parto, idade ao abate, dentre outras. Nestes casos deve-se atentar a estes erros, pois valores nulos para covariáveis causam erros em programas de avaliação genética.

O sétimo erro confere animais que aparecem com número inferior ao dos pais após recodificação executada pelo programa, estes erros são corrigidos automaticamente pelo programa e o número de correções é informado.

O oitavo e último tipo de erro informa os animais e linhas que aparecem mais de uma vez no pedigree, caso o usuário possua um arquivo pequeno e queira recodificar o pedigree direto do banco de dados. Este tipo de erro deve ser desconsiderado em modelos de repetibilidade, análises multicaracterísticas ou modelos de regressão aleatória.

O arquivo de correspondência (CORRESP) é gerado pelo programa RENPED durante a execução da opção um e apresenta uma coluna com o valor codificado do animal seguido da coluna com a identificação real do animal. Este arquivo é utilizado nas opções dois e três do programa RENPED para recodificar os dados e decodificar resultados, respectivamente.

O arquivo de pedigree recodificado (PED-RECOD) é reordenado pela variável animal e reposiciona as colunas animal, pai e mãe, como primeira, segunda e terceira colunas respectivamente, seguindo a ordem original nas demais colunas.

O arquivo de dados recodificados (DATA-RECOD) é semelhante ao arquivo de dados fornecido, entretanto as informações de animal, pai e mãe são recodificadas. O arquivo de resultados decodificado (RES-DECOD) é semelhante ao arquivo de resultados fornecido, entretanto as informações de animal, pai e mãe são as originais.

O arquivo de valores genéticos integrados (INTEGRATED) contém as informações de animal e valor genético para o período integrado, ou seja, para a área abaixo da curva dos valores genéticos de cada animal, integrada para o período solicitado pelo usuário. Este arquivo é fornecido ao executar corretamente a opção quatro do RENPED. O resultado deste arquivo é muito importante para auxiliar a seleção de animais para características produtivas longitudinais, tais como produção de leite no dia do controle e produção semanal de ovos.

Exemplo Aplicado

Para realizar o exemplo foram utilizados 50.398 registros da produção de leite no dia do controle de 2.946 lactações de 1.093 cabras da raça Alpina do Setor de Caprinocultura da Universidade Federal de Viçosa, com informações, provenientes de um rebanho com um pedigree de 7.800 animais.

ortogonal de Legendre para a regressão média da população; α_{im} and γ_{im} são os coeficientes m das regressões genética aditivas e de ambiente permanente, respectivamente, para a cabra i ; 4-1, 2-1 e 7-1 são os graus do polinômio ortogonal de Legendre; t_{ij} é a avariável controle, referente ao dia do controle da cabra i padronizado entre os intervalos de -1 até 1, de acordo com Kirkpatrick et al. (1990); $\Phi_m(t_{ij})$ é a função polynomial de Legendre para o parâmetro m avaliado para o controle j ; e ε_{ij} consiste no efeito específico de cada observação não explicado pelos fatores de correção ou regressões consideradas no modelo.

Os resultados dos valores preditos para os parâmetros das regressões genética aditivas de cada animal gerados pelo WOMBAT no arquivo de soluções foram utilizados para executar a opção quatro do RENPED, opção esta que calcula a área abaixo da curva genética aditiva de cada indivíduo da população. Neste exemplo o RENPED gastou cerca de meio segundo para calcular e mostrar o resultado do valor integrado abaixo da curva levando em conta os valores genéticos preditos dos 6 a 290 dias de lactação de 1046 animais. Parte da agilidade do processo pode estar relacionada ao polinômio ortogonal de Legendre considerado nesta análise para a parte genética aditiva, visto que o polinômio era de ordem dois.

```
CALCULATE THE AREA BELOW OF THE GENETIC CURVE <RRM-LEGENDRE>
OPTIONS:
    1-ORGANIZED FILE;
    2-RnSoln of WOMBAT.2
What is the maximum value for the control variable? 290
What is the minimum value for the control variable? 6

Open the file INTEGRATED.txt to obtain the results!
Pressione qualquer tecla para continuar. . . .
```

Figura 3: Visão do exemplo de utilização da opção cinco do software RENPED.

Posteriormente a realização da integral dos valores genéticos os dados foram submetidos à opção três do RENPED, opção de descodificação dos dados para associar os valores genéticos aos nomes ou códigos originais dos animais. A operação para 1.046 animais com valor genético, em relação a um total de 7.800 codificados no pedigree, gastou aproximadamente um segundo para ser realizada após registrar as informações do arquivo a ser avaliado no programa e acionar seu funcionamento.

Considerações finais

O RENPED é uma alternativa para corrigir erros e recodificar dados alfa numéricos, assim como para viabilizar os arquivos de pedigree e dados para análises em diversos programas, tais como ENDOG, WOMBAT, MTDFREML, DFREML, REMLF90.

O programa RENPED foi testado apresentando resultados confiáveis e rápidos de recodificação e descodificação de arquivos de pedigree, bancos de dados e de resultados de avaliações.

O RENPED foi desenvolvido para funcionar nos sistemas operacionais da Microsoft Windows, entretanto, o programa foi registrado como software livre e seu código pode ser encontrado em Silva (2011), podendo ser modificado e compilado, inclusive para funcionar em outros sistemas operacionais.

Referências Bibliográficas

- Boldman, K. G., Kriese, L. A., Van Vleck, L. D. et al. A manual for use of MTDFREML: a set of programs to obtain estimates of variances and covariances (DRAFT). Lincoln: Department of Agriculture/Agricultural Research Service, 1993.120p.
- Emik, L. and Terril, C. (1949). Systematic procedures for calculating inbreeding coefficients. **J Hered.** 40:51–55
- Gutiérrez, J. P. & Goyache, F. A. A note on ENDOG: a computer program for analysing pedigree information. **Journal of Animal Breeding and Genetics.**v.122, p.172-176, 2005.
- Henderson, C. R. (1976). "A Simple Method for Computing the Inverse of a Numerator Relationship Matrix Used in Prediction of Breeding Values". **Biometrics** 32 (1): 69–83. doi:10.2307/2529339. JSTOR 2529339.
- Meyer, K. DFREML – A set of programs to estimate variance components under an Individual Animal Model. **Journal of Dairy Science**, v.71 (Supl.2), p.33-34, 1988.
- Meyer, K. WOMBAT - A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). **Journal Zhejiang University Science B**, v. 8 n. 11, p. 815- 821, 2007.
- Quaas, R L. 1976. Computing the diagonal elements and inverse of a large numerator relationship matrix. **Biometrics** 32:949.

Silva, F. G. Estudo da produção de leite de caprinos utilizando modelos de regressão aleatória. Viçosa, UFV, 2011. 143p. Dissertação (Mestrado), Universidade Federal de Viçosa.

Wright, S. (1922). Coefficients of inbreeding and relationship. **The American Naturalist**. 56:330. Pág.: 330-338.

CAPÍTULO 4

Considerações na avaliação genética da produção de leite no dia do controle

Resumo: Não existe um consenso atual sobre o ponto de truncamento mais adequado a ser utilizado para a duração da lactação em avaliações genéticas da produção de leite, nem estudos que comparem se diferentes pontos de truncamento podem ocasionar alterações na estrutura genética de populações. Nas avaliações da produção de leite no dia do controle, faltam estudos que demonstrem vantagens na utilização do efeito fixo ano-estação do controle ao invés do parto, além disso, falta maior discussão quanto a melhor forma de compor este efeito. A utilização de variáveis com muitas casas decimais ou com valores muito altos pode em alguns softwares ocasionar problemas de erros numéricos. Para testar as diferentes considerações supracitadas na avaliação genética da produção de leite no dia do controle foram utilizados 21.610 registros da PLDC de 1.043 cabras da raça Alpina do setor de caprinocultura da Universidade Federal de Viçosa. O software RENPED foi utilizado para corrigir erros de pedigree e recodificar os dados alfanuméricos. O programa WOMBAT foi utilizado nas avaliações genéticas, fornecendo previsões dos valores genéticos e critérios de comparação de modelos LogL, AICm e BICm. Foi encontrada diferença nos valores genéticos e classificações dos animais conforme alterações no ponto de truncamento da lactação, existindo grande diferença entre a avaliação considerando pontos de truncamento curtos em comparação a médios e longos. Melhores resultados são obtidos em modelos que utilizem o efeito fixo ano-estação do controle leiteiro ao invés do parto. Também é vantajosa a construção destes efeitos fixos agrupando estações que comecem em um ano e terminam em outro ao invés de apenas mesclar ano e estação. A utilização da produção de leite no dia do controle em litros não atrapalha as avaliações genéticas através de erros numéricos, mas unidades de medidas muito grandes ou muito pequenas podem ocasionar este tipo de problema.

Palavras-chave: efeito de ano-estação, erros numéricos, ponto de truncamento

Introdução

A característica mais trabalhada nos programas de melhoramento genético de animais de aptidão leiteira é a produção de leite acumulada. Mais recentemente, com o surgimento dos modelos de regressão aleatória, a produção de leite no dia do controle (PLDC) também tem passado a ganhar destaque em avaliações genéticas. Em ambos os casos, diferentes pontos de truncamento vem sendo utilizados para determinar o número de dias máximos a ser considerado na avaliação genética da produção de leite. Desde 305 dias, valor padronizado para bovinos até pontos ligados à média da duração da lactação do rebanho caprino em avaliação.

Sarmiento et al. (2006) consideraram a produção de leite até os 305 dias de lactação, enquanto Pereira et al. (2008), trabalhando com cabras da raça Alpina, utilizaram ponto de truncamento de 270 dias, Sarmiento et al. (2008) trabalhando com modelos de regressão aleatória considerou valores da produção de leite entre a segunda e 39ª semanas (7 à 273 dias), e Menezes et al. (2008), trabalhando com cabras da raça Saanen, consideraram produção até os 268 dias de lactação, condizente com a média de duração da lactação do rebanho.

A escolha de um ponto de truncamento de 305 dias de lactação pode não ser adequada para o estudo da produção de leite acumulada de caprinos por não condizer com a fisiologia destes animais. Assim como considerar a média de dias em lactação como limite superior do calculo da PLAC pode desfavorecer os animais com maior persistência de lactação.

Menezes et al. (2007) compararam produções acumuladas até diferentes dias de lactação (200, 220, 240, 260, 280, 305, 340 e 365), e encontraram maior herdabilidade para produção acumulada até os 240 dias. Este é um critério mais interessante do que os anteriores para se definir o ponto de truncamento, pois traz a vantagem de selecionar o período que apresenta maior proporção da variação genética sobre a variação total, no qual se espera maiores ganhos genéticos. Entretanto, esta escolha também não leva em consideração a análise prévia do sistema de produção, o qual pode exigir animais de menor duração de lactação, com maiores picos de produção. Também pode ocorrer o contrário, um sistema que necessite de animais mais equilibrados com maior duração da lactação. Neste caso, considerar apenas 240 dias seria como selecionar pelo período parcial, o qual necessitaria de acompanhamento contínuo da produção complementar.

Silva (2011) ressalta que nos casos supracitados, foram escolhidos pontos de truncamento que influenciaram o sistema produtivo, enquanto a prática mais adequada

seria analisar a fisiologia reprodutiva e produtiva das cabras a fim de se determinar um sistema de produção a partir do qual seja possível definir o ponto de truncamento.

Essa prática deveria ser realizada não só para determinação da duração da lactação para avaliações da produção de leite acumulada, mas também para definir os limites da variável controle dias em lactação em modelos de regressão aleatória. Mas ainda são necessários estudos que demonstrem os reais impactos e diferenças em se praticar a seleção utilizando diferentes durações da lactação, e se isto pode impactar no perfil genético do animal selecionado.

Silva et al. (2013a), na definição de fatores que afetam a PLDC, encontraram efeito significativo da interação entre ano e estação, para caprinos das raças Alpina e Saanen. Como solução para esse efeito comum de ser encontrado, normalmente tem sido recomendado a composição de uma variável única ano-estação, mas não são encontrados trabalhos na literatura que demonstrem que a composição desta variável através da mescla de ano e estação traz o inconveniente de agrupar a estação chuvosa do início do ano com a do final do ano, quando seria mais adequado agrupá-la com o final do ano passado, e a estação do final do ano atual com a estação do início do ano seguinte.

Ou seja, janeiro de 2012 é mais parecido com dezembro de 2011 do que com dezembro de 2012, enquanto este último é mais parecido com janeiro de 2013. Isto pode ser realizado utilizando um sistema de condições e tarefas ao invés de simplesmente mesclar variáveis, mas estudos são necessários para comprovar a vantagem desta modificação.

Outro questionamento com relação à utilização da variável ano-estação do parto como efeito fixo na avaliação da PLDC consiste na falta de conhecimento a respeito das possíveis vantagens de se utilizar o ano-estação do controle leiteiro, o que poderia estar relacionado a uma correção mais eficaz.

Na avaliação de caprinos de aptidão leiteira é muito comum utilizar a produção de leite em litros, mas não são encontrados estudos que demonstrem que oscilações na escala de medida poderiam trazer resultados diferentes, e se o padrão utilizado traz alguma desvantagem, como problema no armazenamento de informações ocasionando resultados ruins nas análises, influenciados pela ocorrência de erros numéricos.

Objetivou-se com este trabalho testar e comparar as rotinas supracitadas utilizadas na avaliação genética da PLDC com diferentes alternativas que possam vir a ser consideradas na melhoria da qualidade das avaliações.

Material e Métodos

Foram utilizados dados da PLDC de cabras do setor de caprinocultura da Universidade Federal de Viçosa (Viçosa, MG). Os controles leiteiros foram realizados semanalmente em ordenha mecânica, realizada duas vezes ao dia. A coleta na parte da manhã teve início às seis horas e a coleta na parte da tarde teve início às quatorze horas. A PLDC representa a soma do leite coletado nestes dois períodos.

Após a análise crítica dos dados restaram 21.610 registros de PLDC da primeira lactação de 1.043 cabras da raça Alpina. O software RENPED (Silva, 2011) foi utilizado para corrigir erros de pedigree e recodificar os dados alfanuméricos. O programa WOMBAT (Meyer, 2007), versão atualizada de setembro de 2013, foi utilizado em todas as análises genéticas, fornecendo as estimativas dos parâmetros e previsões dos valores genéticos para cada animal avaliado. O mesmo programa também forneceu os critérios de comparação de modelos LogL, AICm (-1/2 AIC) e BICm (-1/2 BIC) para as comparações em que se fizeram pertinentes.

Os modelos utilizados nas avaliações da PLDC continham combinações de ajuste para os polinômios das regressões fixa, genética aleatória e de ambiente permanente de ordem 4, 2 e 7, além do número de classes de variâncias residuais igual a 5. Seguindo o melhor modelo encontrado por Silva et al. (2013b), para o mesmo rebanho. O modelo animal unicaracterístico de regressão aleatória utilizado em cada uma das análises é descrito de forma generalizada a seguir:

$$y_{ij} = EF_i + \sum_{m=0}^{4-1} b_m \phi_m(t_{ij}) + \sum_{m=0}^{2-1} \alpha_{im} \phi_m(t_{ij}) + \sum_{m=0}^{7-1} \gamma_{im} \phi_m(t_{ij}) + \varepsilon_{ij}$$

, em que: y_{ij} corresponde à produção de leite no controle j da cabra i; EF_i consiste no conjunto de efeitos fixos, sendo a parte classificatória constituída da média, do grupamento genético (Alpina > 90% = 1; 80% < Alpina ≤ 90% = 2; 70% < Alpina ≤ 80% = 3; 60% < Alpina ≤ 70% = 4; 50% < Alpina ≤ 60% = 5), do ano-estação e tipo de parto. A idade da cabra ao parto foi incluída como covariável, com efeitos linear e quadrático; b_m é o coeficiente de regressão m da PLDC sobre o polinômio de Legendre para modelar a curva média da população; α_{im} e γ_{im} são os m coeficientes de regressão genéticos aditivos e de ambiente permanente, respectivamente, para a cabra i; 4-1, 2-1 e 7-1 são os graus dos polinômios ortogonais de Legendre; t_{ij} é a variável controle, semana de

lactação da cabra i padronizada para o intervalo de -1 a 1; $\Phi_m(t_{ij})$ é a função polinomial de Legendre para o parâmetro m avaliada para a idade t_{ij} ; e ε_{ij} consiste nos efeitos específicos à cada observação não explicados pelos fatores de correção nem pelas regressões do modelo.

Para estudar e entender divergências entre os diferentes efeitos fixos, durações das lactações e unidades de medidas foram realizadas análises, de todos os arquivos que serão mencionados adiante, utilizando o mesmo modelo, caso contrário poderia haver confundimento entre as divergências testadas e o efeito de modelagem.

Para discutir sobre a composição da variável ano-estação foram utilizadas quatro análises, cada uma com um arquivo de dados com diferente composição do efeito fixo ano-estação. O arquivo AEPM.txt, continha o efeito fixo AEPM que consiste em uma variável que mescla os valores de ano e estação do parto, o arquivo AECM.txt continha o efeito fixo AECM que consiste em uma variável que mescla os valores de ano e estação do controle leiteiro. Os arquivos AEPI.txt e AEI.txt, são respectivamente semelhantes aos anteriores, à diferença está na criação das suas variáveis que agrupa a estação do início do ano atual com a estação do final do ano anterior, e a do final do ano atual com a estação do início do ano seguinte.

Para discutir sobre utilização de diferentes durações de lactação foram criados nove arquivos, do PL210.txt ao PL290.txt, variando de dez em dez. Cada arquivo foi modificado em relação ao banco de dados original para conter apenas produções de leite dos seis dias de lactação até o número de dias especificado no arquivo.

Para discutir sobre a utilização de diferentes unidades de medida e possíveis problemas de erros numéricos associadas às mesmas, foram utilizados os arquivos: mililitro.txt, centilitro.txt, decilitro.txt, litro.txt, decalitre.txt, hectolitro.txt, kilolitro.txt, decakilolitro.txt e hectokilolitro.txt. Unidades de medidas superiores ou inferiores a estas não permitiram convergência do software de avaliação genética.

Os resultados dos valores genéticos para os parâmetros do modelo aditivo foram integrados, utilizando o RENPED, para se obter o valor genético acumulado de cada animal em cada lactação. O mesmo programa também foi usado para calcular as correlações de Pearson e Spearman entre os valores genéticos preditos acumulados para a lactação de cada cabra, comparando as diferentes considerações durante o processo de análise.

Resultados e Discussão

Segundo Ribeiro (1997), no hemisfério sul, animais criados em locais com latitudes menores que -20° apresentam sazonalidade reprodutiva influenciada pelo fotoperíodo, enquanto animais criados em latitudes maiores que -20° não apresentam este problema, ocorrendo cio durante todo ano, o que permite a utilização do sistema de criação alternativo, com dois partos ao ano sem a utilização de programas de iluminação artificial complementar, e sem a necessidade de se aplicar hormônios que regulam o cio.

Silva (2011) descreve dois sistemas hipotéticos de produção de caprinos de aptidão leiteira, um que respeita a sazonalidade reprodutiva destes animais, buscando um parto por ano, chamado de sistema tradicional (ST). O outro sistema, denominado alternativo (SA), abordado pelo autor considera a possibilidade de realizar três partos em dois anos, contornando o problema de sazonalidade reprodutiva destes animais através da utilização de programas de luz artificial complementar e aplicação de hormônios indutores de cio.

Animais criados nas condições do sistema tradicional devem apresentar boa genética associada à persistência da lactação, enquanto no sistema alternativo, maiores picos de produção são importantes em detrimento da persistência, buscando-se maior produção em três picos de lactação em torno de dois anos.

Apesar da pertinente descrição feita pelo autor, a respeito de possíveis alterações genéticas em populações sob seleção, este não realiza nenhum tipo de análise que comprove que a seleção da população considerando maiores lactações, com um parto por ano, possa levar resultados distintos da seleção considerando lactações mais curtas, com aproximadamente três partos em dois anos.

A Tabela 1 apresenta os resultados das correlações de Pearson (acima da diagonal principal) e Spearman (abaixo da diagonal principal) para os valores genéticos preditos acumulados para os animais fenotipados.

Os resultados encontrados demonstram que quando se utiliza modelos de regressão aleatória utilizando POL de baixa ordem para a curva genética aditiva, na avaliação da produção de leite no dia do controle, existe uma inversão muito grande no resultado da avaliação de acordo com o número de dias em lactação.

Estes resultados corroboram com a sugestão de se considerar a biologia dos animais e definir qual o sistema de produção será utilizado antes de começar a praticar a seleção dos animais, possivelmente selecionar animais para produção de leite até os 290 dias exigem muito mais genes de persistência em relação a seleção de animais de lactação

mais curtas e maiores picos de produção, a qual espera-se que esteja associada a outros genes.

Se selecionar animais de curta duração da lactação traz alterações muito distintas em relação à animais de longa duração da lactação, por outro lado a seleção a partir de 240 dias de lactação até 290 não trará grandes alterações na estrutura genética da população sob seleção. O que indica que a maioria dos critérios mencionados anteriormente, para escolher o melhor ponto de truncamento da duração da lactação, produziriam resultados semelhantes no processo de seleção genética.

Tabela 1: Correlações de Pearson, acima da diagonal principal, e Spearman, abaixo da diagonal principal, para os valores genéticos preditos acumulados para as cabras, realizando difentes avaliações genéticas considerando pontos de truncamento de 210 a 290 dias nas análises.

| | 210 | 220 | 230 | 240 | 250 | 260 | 270 | 280 | 290 |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 210 | - | 0,9861 | 1,0000 | 0,0934 | 0,0884 | 0,0821 | 0,0841 | 0,0858 | 0,0847 |
| 220 | 0,9832 | - | 0,9861 | 0,0911 | 0,0866 | 0,0812 | 0,0842 | 0,0851 | 0,0826 |
| 230 | 1,0000 | 0,9832 | - | 0,0934 | 0,0884 | 0,0821 | 0,0841 | 0,0858 | 0,0847 |
| 240 | 0,1217 | 0,1169 | 0,1217 | - | 0,9952 | 0,9891 | 0,9852 | 0,9721 | 0,9661 |
| 250 | 0,1146 | 0,1100 | 0,1146 | 0,9946 | - | 0,9946 | 0,9901 | 0,9766 | 0,9708 |
| 260 | 0,1082 | 0,1050 | 0,1082 | 0,9884 | 0,9938 | - | 0,9964 | 0,982 | 0,9768 |
| 270 | 0,1074 | 0,1058 | 0,1074 | 0,9819 | 0,9873 | 0,9948 | - | 0,9889 | 0,9838 |
| 280 | 0,1058 | 0,1033 | 0,1058 | 0,9651 | 0,9714 | 0,9782 | 0,9875 | - | 0,9930 |
| 290 | 0,1014 | 0,0978 | 0,1014 | 0,9591 | 0,9658 | 0,9736 | 0,9827 | 0,9933 | - |

Os valores obtidos para critérios de comparação de modelos, tais como Log L e AIC, aumentaram em função do número de dias em lactação considerados na avaliação genética, o que ocorre possivelmente pelo aumento no volume de dados. Dessa forma, tais critérios não permitem neste caso comparar os pontos de truncamento pois são diferentes conjuntos de dados, e por isso não serão apresentados.

A Tabela 2 apresenta os resultados das correlações de Pearson, acima da diagonal principal, e Spearman, abaixo da diagonal principal, entre os valores genéticos preditos utilizando quatro diferentes efeitos fixos (AECI, AECM, AEPI e AEPM).

Tabela 2: Correlações de Pearson e Spearman, acima e abaixo da diagonal principal, respectivamente, entre os valores genéticos aditivos estimados para AEPM, AECI, AEPI e AEPM.

| | AECI | AECM | AEPI | AEPM |
|------|--------|--------|--------|--------|
| AECI | - | 97,27% | 87,12% | 61,22% |
| AECM | 97,18% | - | 86,71% | 61,31% |
| AEPI | 87,05% | 85,85% | - | 61,54% |
| AEPM | 67,55% | 67,89% | 70,79% | - |

AECI agrupa a estação do controle do início do ano atual com o final do ano anterior; AECM mescla os valores de ano e estação do controle; AEPI agrupa a estação do parto do início do ano atual com o final do ano anterior; AEPM mescla os valores de ano e estação de parto.

Os resultados demonstram que a utilização de avaliações considerando a variável ano-estação como a mesclagem das variáveis ano e estação encontram resultados distintos de quando é considerado o padrão de proximidade entre as estações de diferentes anos. Os resultados mostram divergências ainda maiores quando ao invés de se utilizar o efeito de ano-estação do parto isso é feito em relação ao dia do controle leiteiro.

A tabela 3 apresenta os resultados dos critérios de comparação de modelos LogL, AICm, AICcm e BICm para os quatro modelos testados considerando os quatro tipos de efeitos fixos para corrigir para o efeito de ano-estação.

De acordo com os resultados existe vantagem em agrupar os anos e estações levando em consideração quais são os períodos mais próximos ao invés de desconsiderar e simplesmente agrupar as variáveis. Outro resultado interessante é que se o efeito fixo for criado para a combinação ano-estação incidente sobre o controle leiteiro ao invés do parto, isto também trará vantagens na avaliação da PLDC.

Tabela 3: Critérios de comparações LogL, AICm, AICcm e BICm para os quatro modelos testados.

| Modelo | LogL | AICm | AICcm | BICm |
|--------|---------|---------|---------|---------|
| AECI | 4439,74 | 4403,74 | 4403,68 | 4259,72 |
| AECM | 4427,17 | 4391,17 | 4391,11 | 4247,14 |
| AEPI | 4432,73 | 4396,73 | 4396,67 | 4252,70 |
| AEPM | 4156,60 | 4120,60 | 4120,54 | 3976,57 |

AECI agrupa a estação do controle do início do ano atual com o final do ano anterior; AECM mescla os valores de ano e estação do controle; AEPI agrupa a estação do parto do início do ano atual com o final do ano anterior; AEPM mescla os valores de ano e estação de parto.

A Tabela 4 apresenta os resultados das correlações entre os valores genéticos dos indivíduos preditos utilizando diferentes unidades de medidas, acima da diagonal principal estão expostos valores das correlações de Pearson enquanto abaixo são expostos valores das correlações de Spearman. Os resultados para componentes de variância e critérios de comparação de modelos não foram informados, pois não faz sentido contrastá-los uma vez que são sensíveis a quaisquer alterações nas unidades de medidas.

Os resultados encontrados indicam que não ocorrem erros de predição influenciados por erros numéricos ocasionados por problemas de armazenamento de valores nas variáveis, quando a PLDC é expressa em litros. Entretanto se utilizarmos mililitros, centilitros, decahectalitros ou decaquilolitros, obteremos resultados com oscilações prejudiciais nas estimativas dos valores genéticos e ordenação dos melhores animais com base no valor genético predito.

Tabela 4: Correlações de Pearson e Spearman, acima e abaixo da diagonal principal respectivamente, entre os valores genéticos preditos para os animais, com relação a característica PLDC, comparando diferentes unidades de medida utilizadas na avaliação genética.

| | mL | cL | dL | Litro | daL | hL | kL | dahL | dakL |
|-------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| mL | - | 0,8014 | 0,7420 | 0,7418 | 0,7416 | 0,7424 | 0,7354 | 0,5637 | 0,4857 |
| cL | 0,7952 | - | 0,6628 | 0,6626 | 0,6624 | 0,6636 | 0,6571 | 0,6646 | 0,5656 |
| dL | 0,7359 | 0,6540 | - | 1,0000 | 1,0000 | 1,0000 | 0,9999 | 0,8209 | 0,8025 |
| Litro | 0,7357 | 0,6539 | 1,0000 | - | 1,0000 | 1,0000 | 0,9999 | 0,8209 | 0,8024 |
| daL | 0,7356 | 0,6537 | 1,0000 | 1,0000 | - | 1,0000 | 0,9999 | 0,8208 | 0,8024 |
| hL | 0,7363 | 0,6549 | 1,0000 | 1,0000 | 1,0000 | - | 0,9999 | 0,8219 | 0,8029 |
| kL | 0,7294 | 0,6484 | 0,9999 | 0,9999 | 0,9999 | 0,9998 | - | 0,8197 | 0,8024 |
| dahL | 0,5509 | 0,6506 | 0,8184 | 0,8183 | 0,8182 | 0,8195 | 0,8168 | - | 0,8263 |
| dakL | 0,4611 | 0,5332 | 0,7738 | 0,7738 | 0,7744 | 0,7744 | 0,7737 | 0,8337 | - |

Conclusões

A definição do número de dias em lactação a ser utilizado em avaliações genéticas deve ser realizada no início de um programa de melhoramento. Existindo grande diferença na classificação dos animais quando o ponto de truncamento é muito baixo em relação a pontos de truncamento médios a altos.

Melhores resultados são obtidos em modelos que utilizem o efeito fixo classificatório ano-estação do controle leiteiro ao invés de ano-estação do parto. Melhores resultados também são obtidos através da construção destes efeitos fixos agrupando estações que comecem em um ano e terminem em outro ao invés de apenas mesclar as variáveis ano e estação.

A utilização da produção de leite no dia do controle em litros como tradicionalmente utilizado não atrapalha as avaliações genéticas através de problemas de erros numéricos, mas unidades de medidas muito grandes ou muito pequenas podem ocasionar este tipo de problema.

Referências Bibliográficas

- MENEZES, G. R. O.; MELO, A. L. P.; BRITO, L. F.; ANGELINI, M. S.; COSTA, E. V.; SILVA, F. G.; SARMENTO, R. J. L.; TORRES, R. A.; RODRIGUES, M. T.; LOPES, P. S. **Avaliação de diferentes medidas da persistência da lactação de cabras da raça Saanen utilizando modelo de regressão aleatória.** VII Simpósio Brasileiro de Melhoramento Animal. São Carlos, SP, 10 e 11 de julho de 2008.
- MENEZES, G. R. O.; MELO, A. L. P.; SILVA, F. G.; SILVA, M. F.; TORRES, R. A.; RODRIGUES, M. T. **Avaliação de diferentes medidas da persistência da lactação de cabras da raça Saanen utilizando modelo de regressão aleatória.**

44ª Reunião Anual da Sociedade Brasileira de Zootecnia. Jaboticabal-SP, 24 a 27 de julho de 2007.

MEYER, K. WOMBAT – A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). *Journal Zhejiang University Science B*, v. 8 n. 11, p. 815-821, 2007.

PEREIRA, R. J.; SANTANA JÚNIOR, M. L.; MELO, A. L. P.; MENEZES, G. R. O.; COSTA, E. V.; SILVA, F. G., BRITO, L. F.; ANGELINI, M. S.; TORRES, R. A.; RODRIGUES, M. T. **Fatores ambientais que influenciam características produtivas de cabras da raça Pardo Alpina.** VII Simpósio Brasileiro de Melhoramento Animal. São Carlos, SP, 10 e 11 de julho de 2008.

RIBEIRO, S. D. A. *Caprinocultura: criação racional de caprinos.* Editora Nobel, São Paulo, SP. 1997. p.318. ISBN: 9788521309727.

SARMENTO, J. L. R.; REIS FILHO, J. C.; ALBUQUERQUE, L. G.; LOPES, P. S.; RODRIGUES, M. T. Avaliação genética de caprinos usando a produção de leite no dia de controle. **Revista Brasileira de Zootecnia**, v.35, n.2, p.443-451, 2006.

SARMENTO, J. L. R.; ALBUQUERQUE, L. G.; TORRES, R. A.; RODRIGUÊS, M. T.; LOPES, P. S.; REIS FILHO, J. C. Comparação de modelos de regressão aleatória para estimação de parâmetros genéticos em caprinos leiteiros. *R. Bras. Zootec.*, v.37, n.10, p.1788-1796, 2008.

SILVA, F. G.; BRITO, L. F.; TORRES, R. A.; RIBEIRO JÚNIOR, J.I.; OLIVEIRA, H. R.; CAETANO, G.C.; RODRIGUES, M. T. Factors that influence the test day milk yield and composition. **Genetics and Molecular Research**. 12:1522-1532, 2013a.

SILVA, F. G. **Estudo da curva de produção de leite de caprinos utilizando modelos de regressão aleatória.** 2011, 144p. Dissertação (Mestrado em Genética e Melhoramento) – Universidade Federal de Viçosa, Viçosa – MG. 2011.

SILVA, F.G.; TORRES, R.A.; BRITO, L.F.; EUCLYDES, R. F.; MELO, A. L. P.; SOUZA, N. O.; RIBEIRO Jr. J.I. and RODRIGUES M.T. Random regression models using Legendre orthogonal polynomials to evaluate the milk production of Alpine goats. **Genetics and Molecular Research**. v.12(4), p.6502-6511, 2013b.

CAPÍTULO 5

Modelos de regressão aleatória utilizando polinômios ortogonais de Legendre na avaliação da produção de leite no dia do controle de cabras da raça Saanen

Resumo: Objetivou-se com este trabalho, comparar modelos de regressão aleatória (MRA), utilizando polinômios ortogonais de Legendre (POL), para avaliação de produção de leite no dia do controle. Foram utilizados 14.511 registros de produção de leite provenientes de 512 cabras da raça Saanen. Os principais critérios de comparação foram máximo para o logaritmo da função de verossimilhança e teste da razão de verossimilhança. Os modelos submetidos à avaliação continham diferentes combinações nas ordens dos polinômios fixo (três até cinco), genético aditivo (dois até oito) e permanente individual (dois até oito) e o número de classes de variância residual foi alterado de (um até seis). O software WOMBAT foi utilizado para todas as avaliações. O modelo de regressão aleatória que continha regressão fixa de ordem três, genética aditiva de ordem dois, permanente individual de ordem oito e cinco classes de variância residual foi considerado o modelo mais parcimonioso dentre os estatisticamente equivalentes ao completo. Modelos utilizando ordens menores e combinações mais próximas da regressão genética aditiva e permanente individual são amplamente testados. Os resultados encontrados mostram que modelos parcimoniosos que utilizem altos graus na parte genética aditiva ou permanente individual apresentam resultados muito semelhantes aos que utilizam altas ordens para ambas.

Palavras-chaves: comparação de modelos, curva de lactação, valor genético

Introdução

Modelos de Regressão Aleatória (MRA), associados à metodologia de modelos mistos, foram inicialmente propostos por Henderson Jr. (1982). De acordo com Araújo et al. (2006), Ptack e Schaeffer (1993) foram os primeiros a utilizarem estes modelos em avaliações da produção de leite no dia do controle, utilizando dados do Quebec Dairy Herd Analysis Service. Tais modelos trazem consigo a vantagem de permitir avaliar animais com poucas medições em controles leiteiros, permitem prever o efeito permanente individual, permitem selecionar para aumento de produção e persistência, predição de valores genéticos mais acurados, etc. Mais vantagens com relação ao uso de MRA podem ser encontradas em Meyer et al. (1989) e Ptak & Schaeffer (1993).

Muitos estudos têm sido realizados para comparar e selecionar modelos de regressão aleatória para avaliar geneticamente animais com características produtivas do tipo longitudinal, principalmente em rebanhos de aptidão e foco em produção de leite. Em caprinos de aptidão leiteira da raça Saanen, Menezes et al. (2008) testaram diferentes modificações de modelos de Wilmink, destas, as duas melhores foram comparadas por Menezes et al. (2010) com quatro modelos utilizando polinômios ortogonais de Legendre (POL), sendo encontrado melhor resultado para um POL, o qual continha ordem três para a regressão fixa e quatro para as regressões genética aditiva e permanente individual, além de considerar seis classes de variância residual.

MRA utilizando POL são comumente testados em combinações iguais das curvas genética aditiva e permanente individual. Entretanto, Silva et al. (2013b) compararam distintas combinações para as regressões fixas de ordem dois até cinco, genética aditiva de um a sete, permanente individual de um a sete e utilizando de uma a seis classes de variância residual (exceto três), totalizando 980 análises. Os resultados encontrados pelos autores contradizem os procedimentos geralmente realizados para seleção do melhor modelo. Os modelos mais parcimoniosos que utilizaram altos graus na parte genética aditiva ou permanente individual apresentam resultados muito semelhantes aos que utilizam altas ordens para as duas partes simultaneamente.

Objetivou-se com este estudo comparar e selecionar MRA utilizando diferentes ordens de POL para as regressões fixa, genética aditiva e de ambiente permanente, além de diferentes ordens de variância residual para avaliar geneticamente caprinos da raça Saanen.

Material e Métodos

Os dados foram obtidos no setor de caprinocultura da Universidade Federal de Viçosa (Viçosa, MG). Após edição e verificação de erros o arquivo de dados continha 14.511 registros da produção de leite no dia do controle de 512 primeiras lactações de caprinos da raça Saanen. Os animais foram criados em sistema de estabulação semi-livre e alimentados com mistura de volumoso e concentrado conforme suas exigências nutricionais. A produção de leite foi obtida a partir da soma das produções dos controles leiteiros da manhã e da tarde, que tinham início às 6:00 e às 14:00, respectivamente.

Os modelos testados possuíam combinações de distintas ordens para os polinômios fixo (F3 até F5), genético aditivo (A2 até A8), e permanente individual (P2 até P8), além de classes de variância residual (H1 até H6). O modelo animal unicaracterístico de regressão aleatória é descrito de forma geral a seguir:

$$y_{ij} = EF_i + \sum_{m=0}^{k_b-1} b_m \phi_m(t_{ij}) + \sum_{m=0}^{k_\alpha-1} \alpha_{im} \phi_m(t_{ij}) + \sum_{m=0}^{k_\gamma-1} \gamma_{im} \phi_m(t_{ij}) + \varepsilon_{ij}$$

Em que: y_{ij} representa a produção de leite no dia de controle j , da cabra i ; EF_i é o efeito fixo ajustado, o qual considera o grupamento genético (Saanen > 90 % = 1; 90 % ≥ Saanen > 80 % = 2; 80 % ≥ Saanen > 70 % = 3; 70 % ≥ Saanen > 60 % = 4; 60 % ≥ Saanen > 50 % = 5), ano-estação do parto e tipo de parto (simples, duplo ou triplo). A idade da cabra ao primeiro parto foi incluída como uma covariável de efeito linear e quadrático; b_m é o coeficiente de regressão m da PLDC no polinômio ortogonal de Legendre para modelagem da curva média da população; α_{im} e γ_{im} são os m coeficiente de regressão genético aditivo e permanente individual, respectivamente para a cabra i ; k_b , k_α , k_γ são as ordens dos LOP utilizados; t_{ij} é a variável controle, dia da lactação da cabra i padronizado entre os valores de -1 e 1; e ε_{ij} especifica o efeito para cada observação que não é explicado pelos fatores de correção e regressões do modelo.

Os seguintes critérios foram utilizados para contrastar os modelos estudados: Logaritmo do máximo da função de verossimilhança (LogL), Critério de Informação de Akaike e Bayesiano modificados (AICm e BICm, respectivamente), teste da razão de verossimilhança (TRV), convergência utilizando o algoritmo AI-REML e número de condição (NC). Mais detalhes sobre estes critérios de seleção de modelos podem ser encontrados em Silva et al. (2013b).

O software WOMBAT (Meyer, 2009), foi utilizado em todas as avaliações genéticas, fornecendo estimativas dos parâmetros e valores necessários para calcular os critérios de comparação de modelos.

Resultados e Discussão

Os resultados para o critério de comparação de modelos LogL estão apresentados na Figura 1, os resultados do AICm e BICm não foram apresentados na forma de figura, pois possuem comportamento muito semelhante, não fornecendo informações que possam auxiliar a discussão. Note que na Figura 1, o LogL está organizado em função da ordem das regressões fixa, genética aditiva e permanente individual, além do número de classes de variância residual.

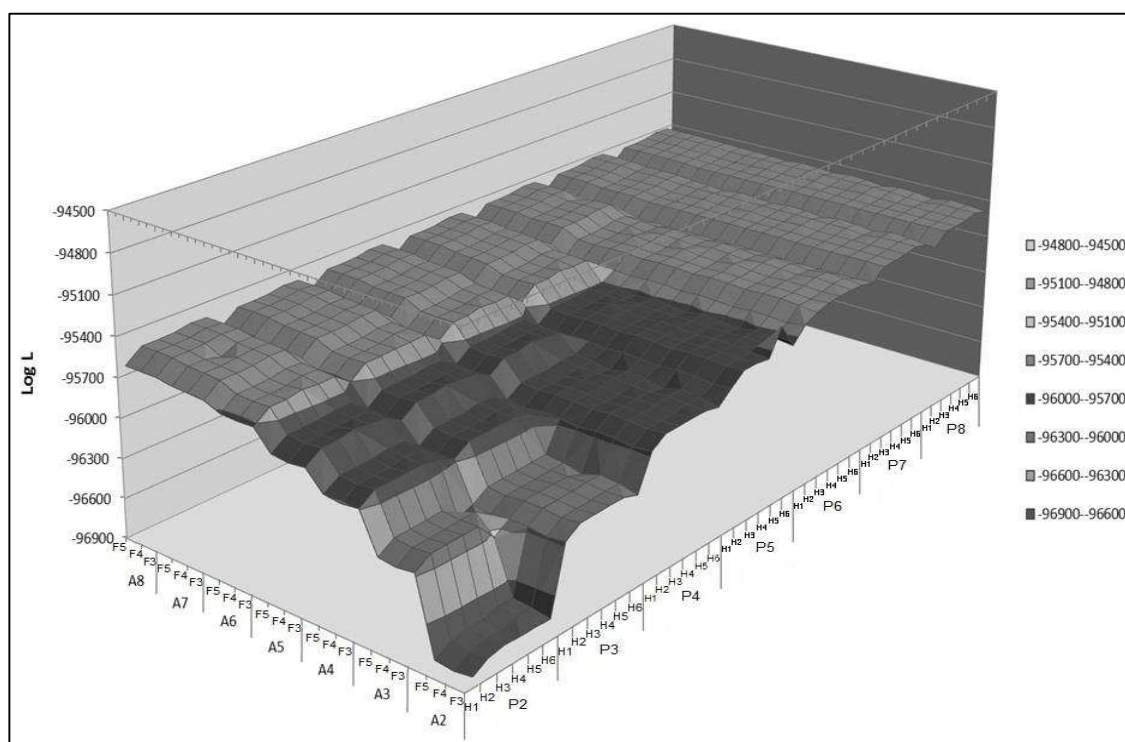


Figura 1: Resultados do logaritmo do máximo da função de verossimilhança em relação ao número de classes de variância residual (H) e ordens dos polinômios genético aditivos (A), permanente individuais (P) e fixos (F), utilizados nas análises.

Os resultados do Teste da Razão de Verossimilhança ($P < 0.01$) são apresentados na Figura 2, o modelo mais parcimonioso estatisticamente igual ao modelo completo (A8P8F5H6). Os resultados do TRV indicaram que 31 modelos dos 882 testados foram estatisticamente iguais ao modelo completo.

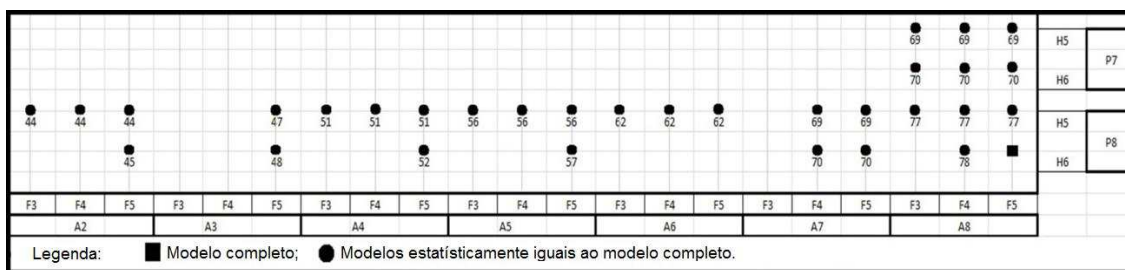


Figura 2: Resultado do teste da razão de verossimilhança, mostrando o modelo completo e os modelos estatisticamente equivalentes.

De acordo com TRV, o modelo mais parcimonioso indicado para ser utilizado em avaliações genéticas do rebanho analisado foi o que considerava regressão fixa de ordem 3, genética aditiva de ordem 2, permanente individual de ordem 8, além de 5 classes de variância residual.

Silva et al. (2013b) recomendam que o melhor modelo de avaliação seja testado para cada rebanho a ser submetido a avaliações genéticas. Neste trabalho foram encontrados resultados semelhantes para o melhor modelo, utilizando o TRV, mas não idêntico ao descritos pelos autores supracitados, que trabalharam com caprinos da raça Alpina. Isso indica que a raça considerada no estudo pode também influenciar o melhor modelo.

Através da Figura 1 é possível perceber que as alterações na ordem da regressão permanente individual e genética aditiva resultam em maior incremento do critério LogL em relação a aumentos da ordem da regressão fixa ou do número de classes de variância residual. Entretanto, à medida que a ordem das regressões genética aditiva e/ou permanente individual ficam mais altas, novos incrementos refletem cada vez menos em aumentos dos critérios utilizados, o que aumenta comparativamente a importância das alterações nas curvas fixas e número de classes de variância residual, não devendo estas serem desconsideradas.

Deste raciocínio também é possível deduzir que futuras comparações de modelos podem ser menos laboriosas se primeiro forem analisadas variações das regressões genética aditiva e permanente individual fixando-se as demais, regressão fixa e as classes de variância residual. Entretanto, após encontrar a melhor combinação das primeiras, não basta considerar variações das segundas, deve-se buscar também variações próximas ao modelo mais parcimonioso encontrado.

Os resultados expressos na Figura 3 mostram variância de ambiente permanente muito alta e variância genética aditiva muito baixa durante toda a lactação, resultados que discordam dos encontrados por Menezes et al. (2010), também trabalhando com a raça Saanen, que encontrou variâncias genética aditivas mais altas no início da lactação e

menores no final, e variâncias de ambiente permanente muito maiores no final da lactação. Estes autores testaram alguns MRA utilizando POL e funções paramétricas de Wilmink, obtiveram como melhor modelo um POL com ordem para regressão genética aditiva igual a quatro e ambiente permanente igual a seis.

Os resultados obtidos neste trabalho também discordaram dos obtidos por Silva et al. (2013) trabalhando com caprinos da raça alpina e utilizando o modelo F4A2P7H5, os quais por sua vez discordaram dos obtidos por Menezes et al. (2010). Estes autores observaram maior variância genética aditiva no final da lactação e variância de ambiente permanente maior no início da lactação. O que segundo os autores pode estar associado a uma maior variabilidade genética associada à persistência da lactação nesta raça. As diferenças entre os três trabalhos também podem estar associadas a um problema de superparametrização de alguns modelos e ou grande diferença entre os graus dos polinômios das regressões genética aditiva e de ambiente permanente. Mais estudos a respeito desta possibilidade ainda são necessários.

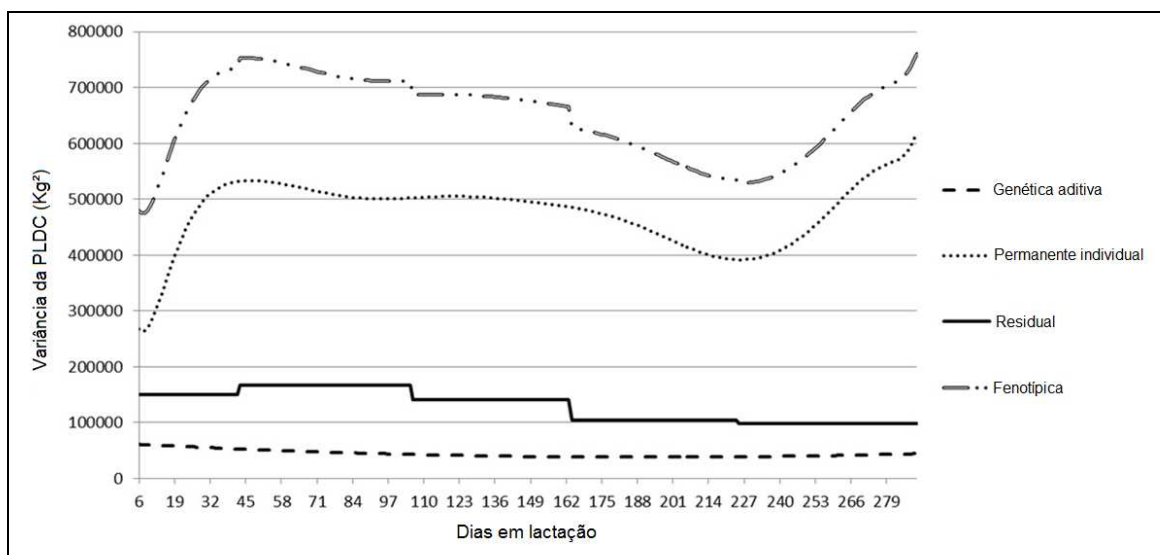


Figure 3: Variâncias fenotípica, genética aditiva, permanente individual e residual para a produção de leite no dia do controle utilizando o MRA F3A2P8H5.

Os autovalores encontrados para o determinante das matrizes de covariância e números de condição de 9,25 e 270,91, para os efeitos genéticos aditivos e permanentes individuais, respectivamente, indicam não haver fortes problemas de multicolinearidade entre os parâmetros do modelo escolhido.

Outra consideração importante que pode ser observada na Figura 1 se refere às variações da regressão genética aditiva ou permanente individual produzirem melhorias nos critérios de comparação tão boas quanto aumentos simultâneos da ordem das duas regressões. Estes resultados corroboram com os encontrados por Silva et al. (2013b), trabalhando com caprinos da raça Alpina. Tais resultados a princípio podem indicar que

modelos de regressão aleatória, apesar de eficientes em explicar variações nos dados, tem dificuldades de separar quais variações seriam de causa permanente individual ou genética aditiva.

Quanto aos algoritmos de estimação de componentes de variância, Misztal (2008) descreve que a convergência utilizando AI-REML é alcançada em apenas algumas rodadas. Entretanto, modelos superparametrizados ou com poucas informações disponíveis para avaliar algum dos componentes de variância podem fazer com que a convergência não seja atingida, ou que estimativas incoerentes ou fora do espaço paramétrico sejam encontradas.

Os resultados encontrados neste trabalho corroboram com a descrição anterior, o AI-REML se mostrou um algoritmo rápido e forneceu estimativas admissíveis. Entretanto, quando os valores de partida utilizados eram de baixa qualidade, comumente o algoritmo caminhava para estimativas fora do espaço paramétrico e não chegava a uma convergência. Este problema era agravado quando modelos com altas ordens para as regressões genética aditiva e de efeito permanente individual eram avaliados. Para alcançar a convergência pelo AI-REML nestes casos, nós utilizamos algoritmos mais lentos como o PX-EM para obter melhores valores de partida e repetimos as análises conseguindo alcançar a convergência utilizando o AI-REML. A maior dificuldade de convergência dentro do espaço paramétrico em alguns modelos mais parametrizados serviu de certa forma como medida da qualidade de convergência das análises.

De fato, a informação mais importante e em média mais precisa sobre a genética de um indivíduo é a que provem do mesmo, e não a que depende de uma informação de parentesco médio. Os baixos valores de herdabilidade ao longo da curva de lactação apresentados na Figura 3 reforçam a teoria de que o confundimento e dificuldade de separação da variação fenotípica em permanente individual e genética aditiva podem ser os responsáveis pelo comportamento gráfico ressaltado a respeito da Figura 1.

Uma crítica também pode ser feita com relação à utilização de ordens distintas para as curvas genética aditiva e de efeito permanente individual, as quais podem fornecer possibilidades de modelagem diferentes a cada um dos efeitos possibilitando vantagens na captação de variação pelo modelo que receba maior ordem. Tal variação nas ordens neste caso de efeitos tão dificilmente separáveis pode causar grande variação em estimativas de componentes de variância e também na predição de valores genéticos.

Para selecionar animais utilizando resultados das avaliações genéticas de MRA é possível utilizar alguns de seus parâmetros que possam estar relacionados ao pico ou persistência da lactação, também é possível escolher dias específicos de produção para

avaliar geneticamente os animais. Outra forma é proposta por Silva et al. (2013c) que consiste em utilizar o software RENPED (SILVA, 2011) para calcular a área abaixo da curva genética aditiva de cada animal, utilizando arquivos organizados ou o próprio arquivo de saída (RnSoln_Animal.dat) do software WOMBAT (MEYER, 2007).

De acordo com Silva et al. (2013b), modelos com ordens maiores para as regressões genética aditivas em relação aos efeitos permanentes individuais apresentam maiores estimativas de herdabilidade ao longo da curva de produção. Segundo os autores, isso não significa que estes modelos devem receber maior ordem de ajuste, por que as estimativas podem também estar erradas, e variações de efeito permanente individual podem estar sendo atribuídas a causas genéticas.

Por outro lado, também pode ocorrer o contrário quando maiores ordens de ajuste são fornecidas para o efeito permanente individual, o que pode ser ainda mais grave na modelagem dos dados e não da biologia a ser considerada, pois de forma geral, o efeito permanente individual tem muito menos pontos e informações por ponto a serem modeladas, sendo mais fácil que este esteja explicando variações nos dados. Ou seja, se a regressão genética aditiva leva também em conta as relações de parentesco e o efeito permanente individual, então possivelmente ordens maiores na modelagem da curva de efeito permanente individual em relação à genética aditiva podem estar captando oscilações nos dados, que deveriam compor a variância residual ou a parte genética aditiva.

Uma versão mais conservadora poderia restringir também que os efeitos genéticos aditivos recebessem ordem superior à curva fixa média da população, uma vez que valores mais altos podem indicar uma não explicação de fenômeno, mas sim do banco de dados que é por amostragem menor quando associado à genética de uma família. Neste caso cuidado deve ser tomado, pois a curva fixa contempla desvios de ordem genética e de efeito permanente individual, isso implica que modelos de ordem maior podem não trazer vantagem. Já para as regressões genética aditiva ou de efeito permanente individual, o aumento da ordem pode trazer vantagens ou em alguns casos também podem ser apenas uma explicação dos dados.

Mais estudos, inclusive utilizando simulação de dados devem ser realizados para reforçar a compreensão sobre as vantagens e problemas na escolha de melhores MRA e avaliação genética de animais utilizando os mesmos. A grande dificuldade para viabilizar estudos de simulação de características quantitativas e longitudinais será entender toda a biologia que coordena a herança e expressão deste tipo de característica. Podendo alguns genes estar mais associados ao pico da lactação, outros à persistência, outros a resistência

a doenças, outros à plasticidade com relação à vantagens nutricionais, etc. São tantas as características que podem influenciar o comportamento de dados longitudinais, que a narrativa de ausência de explicação biológica para modelos muito parametrizados pode não ser válida.

Conclusões

O modelo mais parcimonioso indicado para avaliações genéticas do rebanho analisado foi o que considerava regressão fixa de ordem 3, genética aditiva de ordem 2, permanente individual de ordem 8, além de 5 classes de variância residual.

Os resultados encontrados indicam que a avaliação de modelos apenas com combinações de mesma ordem para a curva fixa e genética aditiva podem não encontrar o melhor modelo de acordo com os critérios de comparação tradicionais.

Mais estudos, inclusive de simulação de dados são necessários para aumentar a relevância desta discussão. A genética que influencia dados longitudinais de características quantitativas pode ser mais complexa que as demais.

Referências Bibliográficas

- AKAIKE, H. **Information theory and an extension of the maximum likelihood principle**. Proceedings of 2nd International Symposium on Information Theory, Budapest, Hungary, 1973. p.267-281.
- ARAÚJO, C. V.; TORRES, R. A.; COSTA, C. N.; TORRES FILHO, R. A.; ARAÚJO, S. I.; LOPES, P. S.; REGAZZI, A. J.; PEREIRA, C. S.; COBUCI, J. A.; SARMENTO, J. L. R. **Uso de funções ortogonais para descrever a produção de leite no dia de controle pro meio de modelos de regressão aleatória**. Revista Brasileira de Zootecnia. v.35, n.3 p.967-974, 2006.
- HENDERSON Jr., C. R. Analysis of covariance in the mixed model: higher-level, Nonhomogeneous and random regression. **Biometrics**, v.38, p.633-40,1982.
- MENEZES, G. R. O.; BRITO, L. F.; MELO, A. L. P.; SILVA, F. G.; ANGELINI, M. S.; COSTA, E. V.; SARMENTO, J. L. R.; TORRES, R. A.; RODRIGUES, M. T.; LOPES, P. S. **Comparação de modelos de regressão aleatória utilizados na avaliação da produção de leite em cabras da raça Saanen**. In: VII Simpósio Brasileiro de Melhoramento Animal. São Carlos - SP. 10 e 11 de junho de 2008.

- MENEZES, G. R. O; CAETANO, G. C.; BRITO, L. F.; SILVA, F. G.; MELO, A. L. P.; TORRES, R. A.; RODRIGUES, M. T. **Estudo da produção de leite em cabras da raça Saanen utilizando modelos de regressão aleatória.** In: VIII Simpósio Brasileiro de Melhoramento Animal - Melhoramento Animal: UMA VISÃO CRÍTICA. Maringá, PR. 01 e 02 de julho de 2010.
- MEYER, K. WOMBAT – A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). *Journal Zhejiang University Science B*, v. 8 n. 11, p. 815-821, 2007.
- MEYER, K.; GRASER, H.U.; HAMMOND, K. Estimates of genetic parameters for first lactation test day production of Australian Black and White cows. **Livestock Production Science**, v.21, n.1, p.177-199, 1989.
- MISZTAL, I. (2008) Reliable computing in estimation of variance components. *Journal of animal breeding and genetics* [0931-2668] vol: 125. Fasc:6 pág: 363-370.
- PTAK, E. and SCHAEFFER, L.R. 1993. Use of test day yields for genetic evaluation of dairy sires and cows. **Livestock Prod. Sci.**, 34(1-2):23-34.
- SILVA, F. G. **Estudo da curva de produção de leite de caprinos utilizando modelos de regressão aleatória.** 2011, 144p. Dissertação (Mestrado em Genética e Melhoramento) – Universidade Federal de Viçosa, Viçosa – MG. 2011.
- SILVA, F.G.; TORRES, R.A.; BRITO, L.F.; EUCLYDES, R. F.; MELO, A. L. P.; SOUZA, N. O.; RIBEIRO Jr. J.I. and RODRIGUES M.T. Random regression models using Legendre orthogonal polynomials to evaluate the milk production of Alpine goats. **Genetics and Molecular Research**. v.12(4), p.6502-6511, 2013b.
- SILVA, F. G.; TORRES, R.A.; BRITO, L. F.; SILVA, L. P.; MENEZES, G.R.O.; BRITO, L. C.; RODRIGUES, M.T.; EUCLYDES, R. F. Genetic evaluation of Alpine goats using different milk control intervals. **Genetics and Molecular Research**. v.12, p.2455 - 2464, 2013c.