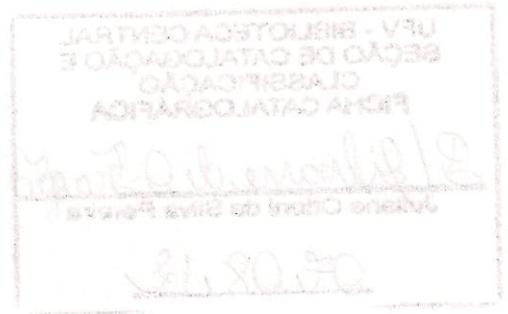


LUIZ FERNANDO BRITO

**MODELAGEM DA PRODUÇÃO DE LEITE DE CABRAS DAS RAÇAS
ALPINA E SAANEN UTILIZANDO REGRESSÃO ALEATÓRIA**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2012



**Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV**

T

B862m
2012

Brito, Luiz Fernando, 1988-

Modelagem da produção de leite de cabras das raças Alpina e Saanen utilizando regressão aleatória / Luiz Fernando Brito. – Viçosa, MG, 2012.

xi, 82f. : il. ; 29cm.

Orientador: Robledo de Almeida Torres.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f. 75-82

1. Caprino - Melhoramento genético. 2. Leite - Produção. 3. Spline, Teoria do. 4. Polinômios. I. Universidade Federal de Viçosa. II. Título.

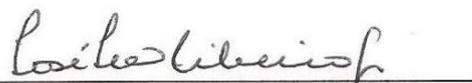
CDD 22. ed. 636.30821

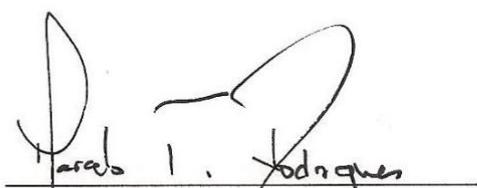
LUIZ FERNANDO BRITO

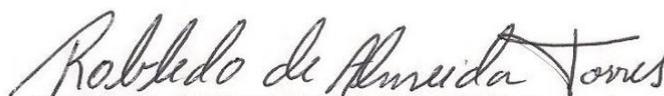
**MODELAGEM DA PRODUÇÃO DE LEITE DE CABRAS DAS RAÇAS
ALPINA E SAANEN UTILIZANDO REGRESSÃO ALEATÓRIA**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do programa de Pós-Graduação em Genética e Melhoramento, para obtenção do título de *Magister Scientiae*.

APROVADA: 16 de julho de 2012


Prof. José Ivo Ribeiro Júnior


Prof. Marcelo Teixeira Rodrigues
(Coorientador)


Prof. Robledo de Almeida Torres
(Orientador)

Aos meus pais Sebastião e Helena,

“Que nos mais simples gestos, me mostraram a importância dos sonhos e dotaram-me de valores essenciais para batalhar por eles com sabedoria e dignidade.”

Dedico este trabalho.

"O mundo está nas mãos daqueles que têm a coragem de sonhar, e correr o risco de viver seus sonhos."

Paulo Coelho

“Após tantos passos dados, ao final de mais um capítulo de minha vida, é impossível não reconhecer os personagens que ilustraram singularmente essa jornada e tornaram esse grande sonho possível: companheiros, camaradas ou simplesmente, pessoas imprescindíveis que chamo de AMIGOS. É eterno o momento de viver e relembrar as alegrias, as histórias e as batalhas que colecionamos ao longo desses anos que estive em Viçosa. Vocês foram presença constante em cada momento. Com certeza, a distância e a ausência não serão sinônimos de esquecimento. As separações são tristes e ao mesmo tempo inevitáveis, mas as lições que aprendi com cada um, carregarei pra sempre e a certeza de que cada reencontro será como se nunca tivéssemos nos separado me conforta muito.

Ainda está nítido em minha memória o dia em que eu saí de casa para vir para Viçosa e despedindo de minha mãe debaixo da velha mangueira, e ela com os olhos cheios de lágrimas me disse: “Oh meu filho, que Deus e Nossa Senhora de Aparecida te iluminem em cada passo que você der e que só coloquem pessoas boas em seu caminho”. Ela não imaginava que aquele moleque franzino encontraria pessoas tão indescritíveis como cada um de vocês. Por isso, chegada a hora de partir, gostaria de deixar meus sinceros agradecimentos a TODOS que tive a honra de conhecer e conviver em Viçosa, pela inestimável lição de vida e por todos os momentos que compartilhamos. Aqui aprendi que mais vale um amigo do que dinheiro no bolso. Sim, aqui aprendi o verdadeiro significado de palavras como amizade, companheirismo e respeito.”

A todos vocês, meus verdadeiros amigos:

Ofereço este trabalho.

AGRADECIMENTOS

As conquistas mais importantes são aquelas que semeamos e colhemos juntos. Hoje sou melhor não só intelectualmente, mas moralmente e isso jamais seria possível sozinho. Por isto só tenho a agradecer:

À Universidade Federal de Viçosa, em especial ao Departamento de Zootecnia e ao Programa de Pós-Graduação em Genética e Melhoramento, pela oportunidade de realização deste curso.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de estudo.

Ao meu orientador, Prof. Robledo de Almeida Torres, pela amizade, conselhos e orientação desde a Iniciação Científica, por ter me apresentado tantas possibilidades ao longo de minha formação acadêmica e por ter me ensinado que sempre vale a pena lutar por nossos sonhos. Foi prazeroso aprender com um profissional tão competente e conviver com um ser humano de tanto caráter!

Ao Prof. Marcelo Teixeira Rodrigues, pelos ensinamentos recebidos desde o início de minha graduação, amizade, confiança, conselhos e acima de tudo por compartilhar essa paixão pela caprinocultura.

Ao Dr. Gilberto Romeiro de Oliveira Menezes, pela amizade, ensinamentos, conselhos, sugestões e principalmente pelo incentivo inicial para seguir no Melhoramento Animal.

À Prof^a. Théa Mírian Medeiros Machado, pelos ensinamentos, conselhos, pela amizade e pelas sugestões.

Ao Prof. José Ivo Ribeiro Júnior, pelos ensinamentos, sugestões e participação na banca examinadora.

A todos os professores e mestres que tive ao longo da vida que me ensinaram a dar passos mais largos e que tanto contribuíram para que eu chegasse até aqui.

Aos meus pais, Sebastião e Helena, exemplos de amor e caráter. Muito obrigado por todo o amor dedicado, pelo esforço para me proporcionar o que há de melhor, pelos valores, pela educação a mim oferecida e por serem meu “porto seguro”. Esta conquista é para vocês!

Às minhas irmãs, Márcia e Mariza, pelo amor, carinho, pela torcida e por mesmo longe, estarem sempre presentes ao longo de todo esse percurso. E

também pela dedicação extra aos nossos pais para amenizar minhas ausências.

Aos amigos do Grupo de Melhoramento Genético de Caprinos (UFV), Ana Lúcia, Edson, Felipe Gomes, Geraldo, Gilberto Menezes, Giovani Caetano, Hinayah, Jô, Laís Brito, Nadson, Paula, Yara, e demais que por descuido eu tenha esquecido. Foi muito bom conviver e aprender com vocês ao longo desses anos.

Aos amigos, Felipe Gomes da Silva e Laís Costa Brito, pela amizade, companheirismo e por não terem medido esforços para me ajudarem na realização deste trabalho.

Aos amigos das repúblicas Al-qaeda e Foca, que sempre foram sinônimos de apoio, companheirismo, alegrias, presença e tudo de mais belo que significa ser amigo.

Aos amigos do GenMelhor, do Melhoramento Animal (“Salinha” e Labtec), da Zootecnia, do Setor de Caprinocultura e demais que conquistei em Viçosa, pela valiosa convivência, compreensão, incentivo e companheirismo desde 2006. Todas as minhas vitórias ficam mais vibrantes e as derrotas mais toleráveis ao lado de vocês.

Aos que acreditam e trabalham pelo sucesso da caprinocultura nacional, obrigado pela parceria! “O entusiasmo, a persistência e a capacidade de se apaixonar por tudo o que fizer serão sempre os segredos da felicidade e da satisfação pessoal e profissional.”

Aos que trouxeram dor, meus sinceros agradecimentos pela oportunidade de crescimento.

A todos os outros que com generosidade e carinho me auxiliaram nesta evolução.

À Deus, pela vida e por todos os motivos que me levam a crer que sou a pessoa mais feliz e realizada desse mundo. Enfim, obrigado pelo que fui, sou e sempre serei: um sonhador.

BIOGRAFIA

LUIZ FERNANDO BRITO, filho de Sebastião Joaquim Brito e Helena Aparecida Sinigalha Brito, nasceu em 07 de novembro de 1988, em Miraí, Minas Gerais.

Cursou o ciclo básico da educação na Escola Municipal Bela Vista (zona rural) e ciclo fundamental e médio na Escola Estadual Santo Antônio, em Miraí, MG.

Em maio de 2006, iniciou o curso de Zootecnia, pelo Departamento de Zootecnia do Centro de Ciências Agrárias da Universidade Federal de Viçosa, em Viçosa – MG, onde foi estagiário voluntário por um ano e meio e bolsista de Iniciação Científica por três anos, acompanhando e desenvolvendo pesquisas na área de Melhoramento Genético Animal, sob orientação do Prof. Robledo de Almeida Torres.

Em julho de 2010 recebeu a Medalha Presidente Bernardes e menção honrosa do Centro de Ciências Agrárias, juntamente com o título de Zootecnista pela Universidade Federal de Viçosa, e em agosto do mesmo ano iniciou o curso de mestrado em Genética e Melhoramento na mesma instituição.

Em setembro de 2010 participou do Curso “Conservación y utilización de recursos Zoogenéticos” em Santa Cruz de La Sierra, Bolívia e em outubro de 2011 participou do XXI Curso Internacional de Melhoramento Animal (INIA), em Madrid, Espanha.

Em 16 de julho de 2012, submeteu-se aos exames finais de defesa de tese para obtenção do título de *Magister Scientiae* em Genética e Melhoramento.

SUMÁRIO

	Página
RESUMO	viii
ABSTRACT	x
1 – INTRODUÇÃO	1
2 – REVISÃO BIBLIOGRÁFICA	4
2.1. A caprinocultura no Brasil.....	4
2.2. Curva de lactação.....	7
2.3. Modelos para análises de dados longitudinais.....	8
2.4. Modelos de Regressão Aleatória (MRA).....	11
2.5. Polinômios segmentados do tipo B (<i>B-splines</i>).....	17
2.6. Heterogeneidade da variância residual.....	24
2.7. Avaliadores da qualidade do ajuste.....	27
3 – MATERIAL E MÉTODOS	29
3.1. Dados.....	29
3.2. Análises.....	30
3.2.1. Efeitos fixos.....	30
3.2.2. Polinômios Ortogonais de Legendre.....	31
3.2.3. Polinômios Segmentados do Tipo B (<i>B-splines</i>).....	33
3.2.4. Comparação dos modelos.....	37
4 – RESULTADOS E DISCUSSÃO	39
4.1. Análise de efeitos fixos (raças Alpina e Saanen).....	39
4.2. Raça Alpina.....	40
4.2.1. Ajuste e comparação de modelos.....	42
4.3. Raça Saanen.....	58
4.3.1. Ajuste e comparação de modelos.....	60
5 – CONCLUSÃO	74
6 – REFERÊNCIAS BIBLIOGRÁFICAS	75

RESUMO

BRITO, Luiz Fernando, M.Sc., Universidade Federal de Viçosa, julho de 2012. **Modelagem da produção de leite de cabras das raças Alpina e Saanen utilizando regressão aleatória.** Orientador: Robledo de Almeida Torres. Coorientadores: Marcelo Teixeira Rodrigues e Théa Mírian Medeiros Machado.

Foram analisados 17.356 registros de produção de leite no dia do controle (PLDC) de 642 primeiras lactações de cabras da raça Alpina e 13.278 registros de 470 cabras da raça Saanen do rebanho da Universidade Federal de Viçosa. O objetivo foi modelar variações na PLDC durante a primeira lactação de cabras das raças Alpina e Saanen através de modelos de regressão aleatória (MRA), usando polinômios ortogonais de Legendre e *B-splines* para a obtenção de modelos mais parcimoniosos e adequados para a estimativa de parâmetros genéticos. As análises foram realizadas usando um MRA unicaracterístico, incluindo os efeitos aleatórios aditivo direto, ambiente permanente e residual. Além disso, o grupo contemporâneo, tipo de parto, agrupamento genético, e os efeitos linear e quadrático da idade da cabra ao parto foram incluídos como efeitos fixos. A curva média de produção de leite, e as funções de covariâncias dos efeitos genético aditivo e de ambiente permanente foram estimados por regressão aleatória utilizando polinômios ortogonais de Legendre de várias ordens (3 a 6) e funções *B-spline* (linear, quadrática e cúbica, com 3 a 6 nós). Além disso, foram avaliados números diferentes de classes de variâncias residuais. Os critérios para seleção dos modelos foram o logaritmo da função de máxima verossimilhança restrita, critério de informação de Akaike, critério de informação Bayesiano e o Teste da Razão de Verossimilhança (TRV). Os componentes de covariância e os parâmetros genéticos foram estimados pelo método da máxima verossimilhança restrita. O programa WOMBAT foi utilizado em todas as análises genéticas. As estimativas de herdabilidade apresentaram tendências semelhantes para ambas as funções e as estimativas foram diferentes para as raças Alpina e Saanen. Os

modelos com um maior número de parâmetros foram mais eficazes para descrever a variação genética de PLDC ao longo da lactação. O MRA mais adequado para avaliação genética da PLDC de caprinos da raça Alpina é uma função *B-spline* quadrática, com seis nós, para a curva média e para as curvas de efeitos genéticos aditivos e de ambiente permanente e cinco classes de variância residual. O MRA mais recomendável para a avaliação genética da PLDC de caprinos da raça Saanen é uma função *B-spline* cúbica, com seis nós e cinco classes de variância residual, para todas as curvas.

ABSTRACT

BRITO, Luiz Fernando, M.Sc., Universidade Federal de Viçosa, July, 2012.
Modeling the milk yield of Alpine and Saanen goat breeds using random regression. Adviser: Robledo de Almeida Torres. Co-advisers: Marcelo Teixeira Rodrigues and Théa Mírian Medeiros Machado.

A total of 17,356 test-day records from 642 first lactations of Alpine goats and 13,278 records of 470 Saanen goats from the herd of Universidade Federal de Viçosa were used in this study. The objective was to model variations in test-day milk yield during the first lactation of Alpine and Saanen goats by random regression model (RRM), using orthogonal Legendre polynomials and B-splines in order to obtain adequate and parsimonious models for the estimation of genetic parameters. The analyses were performed using a single-trait RRM, including the direct additive, permanent environmental and residual random effects. In addition, contemporary group, type of kidding, genetic grouping, kidding order and linear and quadratic effects of the age of goat at calving were included as fixed effects. The mean trend of milk yield, and the additive genetic and permanent environmental covariance functions were estimated by random regression using different orders of orthogonal Legendre polynomial (3 to 6) and on B-spline functions (linear, quadratic and cubic, with 3 to 6 knots) and was evaluated different number of classes of residual variances. The criteria for selection of the models were the logarithm of restricted maximum likelihood function, Akaike information criterion, Bayesian information criterion and the Likelihood Ratio Test (LRT). The covariance components and the genetic parameters were estimated by the restricted maximum likelihood method. The program WOMBAT was used in all genetic analysis. Heritability estimates presented similar trends for the both functions and the estimates were different for Alpine and Saanen breeds. RRM's with a higher number of parameters were more useful to describe the genetic variation of test-day milk yield throughout the lactation. The most suitable RRM for genetic evaluation of PLDC of Alpine goat breed is a quadratic B-spline function, with six knots, to the mean trend and curves

of additive genetic effects, permanent environmental effects and five classes of residual variance. The most recommendable RRM for genetic evaluation of PLDC Saanen goat breed is a cubic B-spline function, with six knots, to all curves and five classes of residual variances.

1 – INTRODUÇÃO

A caprinocultura leiteira é uma atividade potencialmente rentável, em ampla ascensão em todo o mundo e que vem se consolidando, nos últimos anos, como importante segmento da pecuária brasileira. Este grande interesse pela criação de caprinos deve-se, principalmente, à sua versatilidade como produtora de alimentos e matérias primas, tais como leite, carne e pele, bem como à sua alta adaptabilidade às mais diversas condições edafoclimáticas, sendo considerada uma das espécies domésticas mais cosmopolitas. Nesse contexto, apresenta-se como uma atividade promissora e estratégica para um país como o Brasil, que possui vasta diversidade de ecossistemas e realidades socioeconômicas.

Além disso, a crescente demanda por alimentos diferenciados e mais saudáveis também tem despertado grande interesse na exploração da espécie, já que o leite caprino consiste em um alimento de excelente qualidade e de características únicas, como baixo teor de colesterol, menor alergenicidade de sua fração protéica e maior digestibilidade de sua fração lipídica em relação ao leite de vaca.

O Brasil possui cerca de 9,3 milhões de caprinos (IBGE, 2010). Todavia, apesar do elevado efetivo populacional, os índices produtivos e econômicos ainda estão muito aquém dos desejáveis. Para melhorar a rentabilidade da atividade, os produtores têm procurado formas para melhorar o desempenho zootécnico de seus animais, como também reduzir os custos do sistema de produção. E dentre os meios de se alcançar maior eficiência produtiva, um dos mais importantes é o melhoramento genético. Entretanto, pesquisas com caprinos nessa área ainda são escassas. Desse

modo, há necessidade de se realizar trabalhos com raças caprinas de aptidão leiteira em condições edafoclimáticas e de manejo brasileiras a fim de explorar racional e economicamente todo o potencial produtivo e reprodutivo desta espécie e gerar conhecimento para viabilizar programas de melhoramento genético de caprinos no Brasil.

No momento em que o mundo volta sua atenção para as mudanças climáticas, em especial o aquecimento global, a disponibilidade de recurso genético capaz de produzir alimento, em quantidade e qualidade, sob condições adequadas de bem-estar animal, ainda que o ambiente seja adverso e restritivo, os caprinos crescem em importância para a reorganização da pecuária global em bases mais sustentáveis. Para atender às demandas atuais e futuras, os programas de melhoramento devem estar atentos a essas tendências, incorporando novas características e metodologias de avaliação que facilitem o processo de seleção dos animais.

Dentre as características de maior importância para a atividade, destaca-se a produção de leite, que é um exemplo típico de dados longitudinais, caracterizados por observações repetidas em um mesmo indivíduo ao longo do tempo. Para análise destes dados, os Modelos de Regressão Aleatória (MRA) têm sido reconhecidos como mais apropriados, principalmente porque possuem habilidade em ajustar as curvas fixas de lactação para uma população e curvas específicas para cada animal como desvios da curva fixa. Seu uso requer a escolha de uma função para modelar as curvas fixas e aleatórias. Dentre as várias funções, os polinômios ortogonais de Legendre têm sido os mais utilizados atualmente nas avaliações genéticas. Entretanto, o maior problema desses modelos é a

utilização de polinômios de alto grau e a desuniformidade de ajuste ao longo da curva.

Neste contexto, como alternativa aos polinômios de alta ordem, os polinômios segmentados do tipo B (funções *B-splines*) têm sido recomendados, podendo ser encontrados trabalhos em bovinos de leite, e.g. Silvestre et al. (2006), DeGroot et al. (2007), Bohmanova et al. (2008), Bignardi et al. (2011); em bovinos de corte, e.g. Menezes (2010) e Sánchez et al. (2008), em suínos, e.g. Huisman et al. (2002) e em ovinos, e.g. Oliveira (2008). Contudo, não foram encontrados trabalhos, com caprinos, que avaliam a aplicação de MRA utilizando polinômios segmentados do tipo B em avaliações genéticas para dados longitudinais, o que reflete a necessidade de pesquisas buscando os devidos ajustes em função das peculiaridades intrínsecas à espécie.

2 – REVISÃO BIBLIOGRÁFICA

2.1. A caprinocultura no Brasil

A cabra foi o primeiro animal capaz de produzir alimentos, cuja domesticação pelo homem ocorreu há cerca de dez mil anos. Os caprinos têm a mesma origem dos bovinos, com o tronco ancestral dos antílopes e a diferenciação ocorrendo no Plioceno. 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. A cabra doméstica é a *Capra hircus* (RIBEIRO, 1997).

A caprinocultura é uma atividade explorada em todos os continentes, estando presente em áreas que apresentam as mais diversas características edafoclimáticas. 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 baixas produtividade e rentabilidade. Segundo Nogueira Filho e Kasprzykowski (2006), 94% dos caprinos estão nos países em desenvolvimento, sendo uma importante fonte de alimento e renda para essas populações.

No Brasil, os caprinos foram introduzidos pelos colonizadores e se adaptaram bem ao nosso ambiente, principalmente às condições adversas do Semi-Árido Nordestino. Os cruzamentos e os processos de seleção natural e artificial possibilitaram o surgimento de novos tipos raciais com grande rusticidade, embora perdendo um pouco de suas características produtivas. Segundo dados do IBGE (2010), a população caprina no Brasil é

de aproximadamente de 9,3 milhões de animais, com cerca de 90,8 % do efetivo nacional concentrado no nordeste do país. Nesta região, predomina o sistema de criação extensivo em pequenos rebanhos, pertencentes, em sua maioria, a agricultores familiares. Isto evidencia o grande papel social que essa espécie possui sendo uma importante fonte de alimento e renda para as pequenas populações rurais, o que permite que elas se mantenham no campo, reduzindo o êxodo rural.

A criação de caprinos para produção leiteira também tem crescente importância no cenário nacional, sendo que a produção diária de leite chegou a 22.000 litros em 2009 segundo dados do IBGE. Entretanto, de acordo com Bispo (2009), a produção leiteira média dos animais gira em torno de 1 Kg/cabra/dia, sendo que o processo de melhoramento dos rebanhos caprinos não acompanhou a tendência de modernização e intensificação dos sistemas de produção agropecuários, praticando-se, principalmente, uma pecuária pouco tecnificada e composta por animais de genética inferior. A caprinocultura leiteira intensiva prevalece principalmente na região Sudeste, sendo as raças exóticas, Alpina e Saanen, as mais difundidas.

A raça Saanen é originária do Vale Saanen, Suíça e é considerada como a raça leiteira mais difundida em todo o mundo por sua alta produção e persistência na lactação. De acordo com Ribeiro (1997), a raça Saanen possui pelagem uniformemente branca, ou levemente creme, pelos curtos, finos, cerrados, podendo ser um pouco mais longos na linha dorso-lombar e nas partes baixas do corpo (nos machos). Suas orelhas são pequenas ou médias e ligeiramente voltadas para cima, com presença ou não de brincos,

chifres e barba. Sua cabeça é média, cônica, alongada e fina, com a testa bem proporcionada e descarnada. Seu perfil é subcôncavo ou retilíneo. A pele e as mucosas são róseas, podendo as últimas apresentar pequenas manchas escuras, principalmente nas narinas, lábios, mucosa ocular e vulvar, períneo e úbere. Além disso, é uma raça bastante dócil e de fácil manejo em sistemas de criação intensiva.

A raça Alpina é originária dos Alpes Franceses, sendo criada tanto em regiões montanhosas como baixas. Possui pelagem marrom com listra negra no dorso. Assim como a raça Saanen, é altamente especializada para produção leiteira e é bem difundida mundialmente com este propósito.

Embora essas raças estejam presentes há algumas décadas no Brasil, pouco se conhece sobre seu desempenho e são escassas as estimativas de parâmetros genéticos e de outros estudos que dêem suporte a um programa de melhoramento genético no Brasil. De acordo com Gonçalves et al. (2002), no Brasil a melhoria dos plantéis se baseia na importação de animais e sêmen, principalmente da França, onde existe um esquema nacional de avaliação de reprodutores. Dessa forma, a melhoria dos plantéis brasileiros é completamente dependente de outros países, cujos objetivos de seleção nem sempre interessam aos produtores nacionais. E, além disso, pode haver interação genótipo x ambiente, ou seja, os genótipos melhores avaliados na França podem não ter o mesmo desempenho nas diversas regiões brasileiras.

Um exemplo de divergência quanto aos objetivos de seleção é a seleção para produção de leite e constituintes. No Brasil a maior parte do leite produzido é consumida na forma fluida, ao contrário de países como a

França, em que quase a totalidade do leite produzido é utilizada para produção de derivados e por isso é dado um enfoque maior na seleção para porcentagem dos constituintes do leite.

2.2. Curva de lactação

De acordo com Guimarães et al. (2006), na caprinocultura leiteira, a produção econômica do leite é a característica de maior relevância. O conhecimento do perfil e da magnitude da curva de lactação tem importante papel no melhoramento genético de caprinos, além de oferecer subsídios para o planejamento da atividade.

Durante muitos anos, as produções de leite no dia do controle (PLDC's) foram utilizadas apenas para o cálculo da produção total acumulada de cada cabra e essa era utilizada como o principal critério de seleção dos animais. Entretanto, quando se utiliza a produção de leite acumulada desconsidera-se as variações na forma das curvas de lactações entre as cabras, reduzindo a variabilidade genética da produção. No momento da avaliação genética, se a cabra está com a lactação ainda em curso, a produção em 270 dias, por exemplo, é obtida por meio da projeção da lactação parcial baseada nos registros de produção já aferidos, com auxílios de fatores de correção. Assim, cabras que tem maiores persistências, geralmente têm suas produções subestimadas, ao passo que cabras com menores persistências têm suas produções superestimadas. Esta situação pode causar problemas na avaliação genética uma vez que a variabilidade na forma da curva dos animais é desprezada, acarretando eliminação da variação genética existente para produção de leite. Visando

contornar estes problemas, a adoção de modelos que considerem a PLDC vem substituindo a produção acumulada nas avaliações genéticas.

A curva de lactação pode ser definida como a representação gráfica da produção de leite em todo o período da lactação e pode auxiliar na tomada de decisões sobre mudanças no manejo alimentar, descarte de animais ou mesmo em programas de seleção. A sua importância reside na caracterização ampla da produção animal durante toda a lactação, podendo ser identificados: tempo de ascensão ao pico, pico de produção, tempo de queda (persistência de produção ou lactação), duração da lactação, duração do período de serviço, duração da gestação, além de quedas bruscas de produção, respostas a dietas, manejo, etc. (CHAVES, 2009).

Jacopini et al. (2010) também ressaltam que a comparação da forma da curva entre grupos distintos de animais, com diferentes composições raciais, idades ao parto, rebanhos e outros tratamentos de interesse, é de grande importância, pois, mediante essas comparações, podem ser obtidas informações sobre a eficiência desses grupos, propiciando um melhor controle de produção além da seleção daqueles animais mais produtivos.

2.3. Modelos para análise de dados longitudinais

A PLDC é o somatório de todos os registros de produção de leite durante um dia de uma cabra. De acordo com Bignardi (2010), a produção de leite é um exemplo típico de medida repetida ou de dados longitudinais (medidas que são feitas sequencialmente no mesmo animal, ao longo do tempo).

A mensuração repetida de uma característica pode fornecer informações sobre mudanças que nela ocorrem ao longo do tempo. Uma função descrevendo tais mudanças é interessante, pois auxilia a compreensão e a manipulação das transformações da característica.

Características medidas várias vezes ao longo da vida do animal têm sido estudadas de várias formas. A maneira mais simples para analisá-las é assumir que todas as medidas representam a mesma característica e usar um modelo de repetibilidade. Entretanto, neste modelo, pressupõe-se que a correlação entre medidas repetidas seja igual à unidade, portanto, todas as (co) variâncias genéticas e fenotípicas entre as diferentes medidas são de mesma magnitude. Isto é, os mesmos genes controlam o desempenho ao longo do tempo (ALBUQUERQUE, 2004). De acordo com Speidel et al. (2010), na situação em que as observações repetidas seguem algum tipo de curva (e.g., curvas de crescimento ou lactação), correlações entre as observações mais próximas no tempo são maiores do que as mais distantes. Neste caso, um modelo mais complexo que considera a estrutura de correlação diferente entre as sucessivas observações é necessário.

Outra alternativa seria utilizar modelos multicaracterísticos, que admitem cada PLDC como uma característica distinta, além de levarem em consideração as covariâncias entre estas lactações e as informações de todas as lactações na estimação de parâmetros e predição de valores genéticos para cada característica, melhorando a precisão e a acurácia dos mesmos.

Schaeffer (2003) expõe três aspectos que tornam a abordagem multicaracterística mais favorável: i) A correlação genética é muito maior que

a correlação residual; ii) As características têm herdabilidades muito diferentes, sendo que a seleção para a característica de menor herdabilidade é favorecida pelo uso da informação da característica de maior herdabilidade; iii) Nem todos os animais apresentam dados coletados de uma característica de interesse econômico, sendo isto compensado pela análise conjunta com características medidas em todos os animais.

Entretanto, de acordo com Bignardi (2010), a maior desvantagem das análises multicaracterísticas é a dificuldade computacional de analisar um elevado volume de dados e a necessidade de estimar muito mais parâmetros, uma vez que utilizando modelo multicaracterístico completo, cada PLDC é considerada como uma característica distinta, assim o número de efeitos no modelo aumenta com o número de controles considerados.

De acordo com Albuquerque (2004), outra forma de trabalhar com características que se repetem é descrever a trajetória por meio de um modelo matemático (regressão sobre o tempo), estimar os parâmetros e funções da curva e selecionar para os mesmos. Diversos trabalhos de pesquisa têm sido realizados para estimar fatores genéticos e de meio que afetam os parâmetros de curvas de crescimento e lactação e os mesmos podem ser utilizados como critérios de seleção. Entretanto, esta é uma alternativa que raramente tem sido implementada na prática por grandes programas de melhoramento de gado de corte ou de leite.

Outra alternativa para a avaliação da PLDC são os modelos de regressão aleatória. De acordo com Bohmanova et al. (2008), várias funções podem ser usadas para ajustar regressões fixas e aleatórias. As primeiras aplicações usaram funções paramétricas e funções de forma da lactação,

tais como a Ali e Schaeffer (1987) e Wilmink (1987). Entretanto, estas funções vêm sendo substituídas por polinômios ortogonais de Legendre e mais recentemente pelas funções B-splines.

2.4. Modelos de Regressão Aleatória (MRA)

Em 1982, Henderson Jr. propôs uma alternativa para análise de dados longitudinais, que seria o uso de MRA, contornando o problema da super parametrização dos modelos multicaracterísticos.

Os MRA podem ser chamados também de modelos de dimensão infinita, uma vez que existem medidas infinitas por indivíduo e tais medidas são mais correlacionadas, quanto mais próximas entre si (EL FARO e ALBUQUERQUE, 2003). Há vantagens em usar os MRA para avaliações genéticas de características produtivas de gado de leite. O modelo para PLDC considera muito mais fatores ambientais que podem afetar as vacas diferentemente ao longo da lactação. As regressões aleatórias consideram diferentes formas de curva de lactação para cada vaca. Os MRA também permitem que uma vaca possa ser avaliada com base em apenas um registro de PLDC ou qualquer número de registros durante a lactação. Os MRA levam em conta diferentes variâncias genéticas e residuais de produção no decorrer da lactação. Entretanto, a complexidade computacional precisa ser reduzida para tornar os MRA viáveis como procedimentos de rotina (JAMROZIK e SCHAEFFER, 1997).

Em MRA, a regressão fixa descreve a curva média da população, a regressão aleatória genética aditiva descreve a curva de cada animal como diferença da média da população, a regressão aleatória de ambiente

permanente descreve diferenças não aditivas, tais como efeitos de interação intra *locus* e entre *locus* que estejam segregando, além de efeitos ambientais que estão influenciando as medições. As variações não explicadas pela regressão são denominadas como o resíduo (SILVA, 2011).

De acordo com Pereira (2009), os MRA apresentam uma série de vantagens em relação aos modelos tradicionais:

- 1) Não exigem número mínimo de medidas por animal, ao contrário dos modelos que consideram a produção até 305 dias (JAMROZIK e SCHAEFFER, 1997);
- 2) Consideram o parentesco existente entre os animais, ao contrário das funções que descrevem a curva de lactação;
- 3) Não há necessidade de se criarem classes de desempenho arbitrárias ou de utilizar fatores de ajuste para determinada idade (dias em lactação) (ALBUQUERQUE, 2004);
- 4) Permitem a utilização de dados de animais com um ou poucos registros de controles leiteiros (SCHAEFFER e DEKKERS, 1994);
- 5) Permitem a estimação de componentes de covariância entre os controles, incluindo períodos em que os controles não foram realizados (EL FARO e ALBUQUERQUE, 2003);
- 6) As análises abrangem todo o intervalo contínuo em que as medidas foram tomadas, permitindo predizer o valor genético dos animais em qualquer ponto do intervalo em que foram feitas as mensurações fenotípicas;
- 7) Permitem predizer o valor genético dos animais para persistência da produção;

8) Permitem estimar um menor número de parâmetros, principalmente quando se objetiva obter parâmetros genéticos entre várias características, analisadas simultaneamente (REKAYA et al., 1999).

Reis Filho (2009) realizou estudos comparativos das avaliações genéticas para a produção de leite acumulada em 305 dias de lactação (PL305) de animais da raça Gir Leiteiro, por meio de modelos de regressão aleatória e de repetibilidade. De acordo com o autor, utilizando as PL305 das três primeiras lactações, o modelo de regressão aleatória pode ser utilizado como alternativa para incrementar a precisão das avaliações genéticas, proporcionando maiores estimativas de herdabilidade (de 0,26 a 0,28), méritos genéticos preditos semelhantes (correlações acima de 0,97) e alta coincidência entre os animais selecionados (acima de 92%), em comparação com o modelo tradicional de repetibilidade.

Em trabalho com dados simulados de crescimento de bovinos de corte, Meyer (2004) encontrou acurácias, medidas pela correlação entre valores simulados e estimados, de 5,7 a 8,3% maiores para MRA em relação ao modelo multicaracterístico (MM). Segundo a autora, a vantagem se deve à melhor modelagem das variâncias e outros parâmetros genéticos fornecidos no modelo de regressão aleatória. Entretanto, de acordo com Silva (2011), é possível também imaginar que a menor parametrização dos MRA em relação a MM tenha reduzido os problemas de convergência dos algoritmos. Ademais, segundo esse autor, devido à limitada capacidade dos computadores atuais de resolver análises super parametrizadas, o uso de modelos mais parcimoniosos, como os MRA, tornam mais acurados os resultados obtidos.

Como desvantagem dos MRA, vários autores citam o aumento no requerimento computacional, devido ao maior número de efeitos aleatórios por animal, nas equações de modelos mistos; e o fato de que as análises considerando a produção no dia do controle possuem número maior de observações do que as análises considerando a produção em 305 dias (JAMROZIK e SHAEFFER, 1997; SCHAEFFER et al, 2000). Contudo, com o avanço nas tecnologias de *hardware* atuais a preços acessíveis aos programas de melhoramento, este maior requerimento torna-se irrelevante. Portanto, parece aceitável que ênfase maior deva ser dada às propriedades estatísticas e biológicas das metodologias de avaliação genética quando comparadas ao custo computacional das mesmas (PEREIRA, 2009).

A modelagem das curvas de regressão fixa e aleatórias genética aditiva e de ambiente permanente tem sido feita por meio do uso de funções como os polinômios ortogonais de Legendre de várias ordens (BREDA et al., 2006; SARMENTO et al., 2008; FREITAS et al., 2008; BIGNARDI et al., 2011; SILVA, 2011; entre outros). Os polinômios ortogonais de Legendre são funções não paramétricas amplamente utilizadas para estudar a estrutura de covariância para efeitos aleatórios genéticos aditivos e de ambiente permanente em dados longitudinais.

Na maioria dos trabalhos, para modelar dados longitudinais, têm se utilizado regressões sobre polinômios ortogonais de Legendre. Estes modelos não requerem qualquer pressuposição sobre a forma da curva ou trajetória da característica com o tempo.

De acordo com Meyer (2005), os polinômios ortogonais de Legendre têm sido amplamente utilizados. Segundo a autora, polinômios de ordens

mais altas são flexíveis e tem sido capazes de modelar bem as mudanças nas médias e variâncias ao longo de uma escala contínua. Entretanto, esses polinômios frequentemente dão uma alta ênfase nas observações nos extremos e são problemáticos para altas ordens de ajuste. Além disso, polinômios de altas ordens possuem difícil interpretação biológica.

Menezes et al. (2008) compararam o uso da função paramétrica de Wilmink e algumas modificações da mesma, na modelagem das curvas de regressão fixa e aleatória genética aditiva e de ambiente permanente da produção de leite de cabras da raça Saanen. Os cinco modelos comparados diferiam quanto ao termo exponencial da função de Wilmink. O modelo que considerou o menor valor para a constante da parte exponencial da equação foi considerado o melhor, sendo este comparado por Menezes et al. (2010a), com quatro modelos utilizando polinômios ortogonais de Legendre. O melhor modelo foi o que utilizou polinômios ortogonais de Legendre, empregando terceira ordem para a curva fixa, quarta ordem para a curva da variância genética aditiva e sexta ordem para a curva de ambiente permanente, além de considerar seis classes de variância residual.

Cobuci et al. (2004) compararam funções de Wilmink e polinômios ortogonais de Legendre de ordens 3, 4 e 5 na descrição de curvas fixas e aleatórias de lactação, quanto às possíveis alterações nas estimativas de herdabilidade da produção de leite de vacas da raça Holandês. Os autores concluíram que o polinômio ortogonal de Legendre de ordem 4 era o mais adequado, por apresentar, assim como o de ordem 5, menores variações da herdabilidade ao longo da curva de lactação em relação aos demais

modelos, e ser computacionalmente menos exigente do que o polinômio de ordem 5.

Freitas et al. (2008) utilizaram um total de 8.276 observações de PLDC de 1.021 primeiras lactações de vacas da raça Guzerá para contrastarem funções paramétricas (Ali e Schaffer e Wilmink) e polinômios ortogonais de Legendre de terceiro, quarto e quinto grau. Os modelos que incluíram maior número de parâmetros foram os que melhor se ajustaram aos dados. O polinômio de Legendre de quarto grau, por apresentar menor demanda de tempo e não apresentar problemas de convergência na solução do sistema de equações, foi o mais indicado.

Araújo et al. (2006) utilizaram 68.523 controles leiteiros de 8.536 vacas da raça Holandês para contrastarem diferentes funções para descreverem a curva de lactação. As funções utilizadas foram a exponencial de Wilmink, a função de Ali e Schaeffer e os polinômios de Legendre de segundo e quarto graus. Verificou-se que os modelos de regressão aleatória que utilizaram os polinômios de Legendre descreveram melhor a variação genética da produção de leite.

Segundo Meyer (2005), análises de regressão aleatória com polinômios de ordem cúbica, quártica ou mesmo de ordens mais elevadas têm levado a estimativas de componentes de variância e parâmetros genéticos erráticos e implausíveis. Isto tem sido especialmente prevalente em conjuntos de dados que contêm poucos registros nas últimas idades, em conjuntos de dados com poucos registros por animais e em análises em que os animais têm menos registros do que a ordem dos polinômios ajustados.

Silva (2011) analisou 20.710 registros de PLDC de 667 cabras da raça Alpina com o objetivo de identificar o melhor MRA utilizando polinômios ortogonais de Legendre. O MRA utilizando polinômio ortogonal de Legendre mais indicado para avaliação genética da PLDC de caprinos da raça Alpina foi o que considerou a curva fixa de ordem 4, a curva de efeitos genéticos aditivos de ordem 2, a curva de efeitos de ambiente permanente de ordem 7 e ao menos cinco classes de variâncias residuais.

Alternativas para diminuir o grau dos polinômios estão sendo estudadas. Uma delas é a utilização de funções *splines* ou também chamados de polinômios segmentados, que são curvas constituídas por segmentos individuais de polinômios de grau reduzido e que se unem em pontos específicos, denominados nós (De BOOR, 1978).

2.5. Polinômios segmentados do tipo B (B-splines)

Segundo Oliveira (2011), a origem do nome *spline* vem de uma régua elástica, usada em desenhos de engenharia, que pode ser curvada de forma a passar por um dado conjunto de pontos (x_i, y_i) . Os nós eram fixados nas áreas de interesse causando a deformação da estrutura. Apenas na década de 60 sua formulação matemática foi desenvolvida, apesar de ser usada desde o século passado.

De acordo com Paulson (2007), o uso de funções *splines* é interessante quando um modelo polinomial de pequeno grau não ajusta precisamente os dados, e o pesquisador não quer usar uma função polinomial complexa para modelar os dados. Em tais casos, as funções *splines* apresentam-se como excelente opção.

De acordo com Keele (2008), normalmente para os propósitos estatísticos, *splines* são funções de regressão segmentadas em que se impõem restrições a fim de se unirem nos pontos chamados nós. Em suas formas mais simples, *splines* são modelos de regressão com um conjunto de variáveis fictícias no lado direito do modelo que é usado para forçar a linha de regressão mudar de direção em algum ponto na faixa de X. Para a mais simples regressão *spline*, as funções segmentadas são lineares, uma restrição que pode ser alterada. Em essência, se ajusta linhas de regressão separadas dentro das regiões entre os nós, e os nós ligam os ajustes de regressão segmentada. Ou seja, *splines* são um modelo local com ajustes locais entre os nós que permitem estimar a forma funcional dos dados.

De acordo com Varga (2009), as grandes vantagens no uso de uma função polinomial são tornar possível a obtenção de uma curva que passa por todos os pontos conhecidos e a facilidade na obtenção das derivadas e integrais da curva gerada, o que torna os polinômios viáveis na interpolação. Porém, a natureza oscilatória dos polinômios de alto grau restringe o seu uso. Tal fato não ocorre nas *splines* onde a aproximação polinomial de menor grau é feita para cada intervalo do domínio, conseqüentemente há pouca flutuação na curva, além de permitir curvaturas bem diferentes em cada região.

O rápido desenvolvimento das funções *splines* é devido, principalmente, a sua grande utilidade em situações aplicadas. Classes de funções *splines* possuem várias excelentes propriedades estruturais assim como excelentes poderes de aproximação. Como elas são fáceis de armazenar, avaliar e manipular em um computador, uma miríade de

aplicações na solução numérica de uma variedade de problemas em matemática aplicada tem sido desenvolvidas (SCHUMAKER, 2007).

Schenkel (1989) apresentou algumas vantagens das funções *splines*:

- a) Pela inclusão de zeros na coluna das variáveis explanatórias, há uma quebra da multicolinearidade, aumentando a confiabilidade dos valores estimados (y) uma vez que minimiza o fator de inflação da variância (FIV) e o viés dos estimadores;
- b) O comportamento da função que descreve a relação das variáveis em uma região torna-se bastante independente do comportamento em outras regiões com características distintas, o que torna a capacidade de aderência da função *spline* aos pontos superiores a dos polinômios ordinários;
- c) A medida de qualidade do ajustamento, dada normalmente pela soma de quadrados total devido à regressão, será bastante uniforme para a curva inteira do polinômio segmentado, independentemente da uniformidade da distribuição das observações;
- d) Funções não lineares e polinômios ordinários de graus diferentes podem ser aproximados por funções *splines* que são lineares nos parâmetros, de fácil estimação, com propriedades ótimas (*Best Linear Unbiased Estimators* – *BLUE*) e de simples interpretação biológica;
- e) Aumentam o domínio da aproximação, isto é, fornecem estimativas em um campo mais amplo de possíveis valores.

As funções *B-spline* normalmente proporcionam ajustes equivalentes às outras funções *splines*, entretanto de acordo com Meyer (2005), essas têm propriedades numéricas melhores. Segundo a mesma autora, funções *B-splines* englobam um conjunto de funções sobrepostas, suaves e não

negativas, as quais são unimodal e somam a unidade para todos os valores de t , podendo-se ter funções *B-splines* de diferentes graus. De acordo com Rice e Wu (2001) a utilização de *B-splines* é adequada para modelar efeitos aleatórios em análises de modelos mistos, e também eficientes na estimação de funções de covariância.

De acordo com Menezes (2010b), os polinômios segmentados do tipo B apresentam vantagens interessantes em relação aos polinômios ortogonais de Legendre, tal como propriedades numéricas superiores por promover maior esparsialidade e quebra de multicolinearidade. Além disso, o fato de priorizarem o ajuste local em relação ao global os tornam mais robustos em relação a reduzido número de dados para certas faixas de conjuntos de dados.

Para Schenkel (1989), o maior obstáculo para o uso de funções *splines* é que os nós devem ser conhecidos. O conhecimento prévio na área técnica específica, o uso de diagramas e médias para a estimação grosseira de nós ou o uso de um modelo não-linear para a estimação simultânea dos nós e dos coeficientes de regressão podem ser empregados nesta tarefa.

Na prática, o que se tem utilizado nos trabalhos zootécnicos é a fixação dos nós das curvas através da inspeção visual dos pontos num gráfico de dispersão. Para assegurar uma escolha mais correta (que minimize a soma de quadrados dos resíduos) pode-se escolher pontos próximos ao local onde poderia estar situado o nó e então proceder a estimação dos parâmetros do modelo. O ponto adotado seria aquele que minimizasse a soma de quadrados do resíduo. Este método, aparentemente

trabalhoso, é exequível e bastante prático em decorrência da facilidade de computação dos polinômios (SCHENKEL, 1989).

Muñoz-Berrocal et al. (2005), trabalhando com polinômios segmentados para ajuste de curvas de lactação de búfalas da raça Murrah, estimaram os nós da seguinte forma: primeiramente foi feita uma inspeção, utilizando um diagrama de pontos. Segundo os autores, esse procedimento auxilia a visualização de quantos segmentos compõem a reta, além dos prováveis pontos onde ocorrem mudanças na curvatura. Em uma segunda etapa, valores iniciais obtidos através do diagrama foram continuamente experimentados utilizando o procedimento *PROC REG* (SAS,1995), de maneira que se encontrou o ponto onde a Soma de Quadrados do Resíduo foi minimizada.

White et al. (1999) utilizaram uma abordagem não-paramétrica através do uso de funções *splines* cúbicas para o ajuste da curva de lactação. O modelo completo incluiu dez nós e, além deste, os autores utilizaram também modelos reduzidos com três até oito nós localizados na média do dia em lactação de cada controle, comparando os ajustes destes modelos com o obtido usando a curva exponencial de Wilmink. A herdabilidade foi de 0,23 no primeiro controle, 0,20 do quinto ao nono e 0,27 no último controle. Os modelos de *splines* que continham mais que oito nós obtiveram melhor ajuste do que a curva exponencial de Wilmink.

Muñoz-Berrocal et al. (2005) avaliaram o uso de polinômios segmentados, dados pelos modelos quadrático-quadrático e com três segmentos quadráticos, no ajuste da curva de lactação média de búfalos da raça Murrah e seus mestiços. Segundo os autores, os polinômios

segmentados quadrático-quadrático e o polinômio segmentado com três segmentos quadráticos, apresentaram bom ajustamento para a curva média de lactação em búfalas, não havendo diferença significativa de qualidade de ajuste entre ambos os modelos.

Meyer (2005) estimou parâmetros genéticos para pesos do nascimento aos 820 dias de idade em bovinos de corte da raça Angus, aplicando funções B-*splines* para descrever as trajetórias para os efeitos aleatórios de um modelo de regressão aleatória. O autor reportou que este método é menos susceptível a problemas frequentemente observados em análises com regressão polinomial. Esta menor susceptibilidade se deve ao fato do modelo conter segmentos individuais de polinômios de menores graus e um melhor controle da influência global das observações individuais. Um bom ajuste dependerá de uma boa escolha dos graus do polinômio e os locais onde os nós serão inseridos.

Silvestre et al. (2006) estudaram sete diferentes funções matemáticas, dentre elas a de Wood, Wilmink, Ali e Schaeffer, três polinômios de Legendre (ordens 3 a 5) e uma *spline* cúbica para ajuste de dados de produção de leite no dia do controle aferidos do 5^o ao 305^o dia da lactação de bovinos. As lactações foram agrupadas em oito grupos amostrais utilizando-se dois critérios: intervalo do parto ao primeiro controle leiteiro e intervalo de controles leiteiros. Os autores verificaram, em todos os grupos amostrais, a superioridade de ajuste das funções *splines* cúbicas, com 4, 11 ou 32 nós equidistantes, sobre todas as outras funções matemáticas. Tal superioridade ficou ainda mais evidente nos grupos com intervalos do parto ao primeiro controle e intervalo de controles maiores.

Bohmanova et al. (2008), estudando dados de produção de leite, gordura e proteína no dia do controle e também de escores de células somáticas de vacas da raça Holandês no Canadá, compararam quatro modelos de regressão aleatória, em que as regressões fixas e aleatórias eram ambas ajustadas por polinômios de Legendre de ordem 4 ou funções *splines* lineares com quatro, cinco ou seis nós. Foram consideradas 12 classes de variância residual. Os autores concluíram que o modelo que utilizava funções *splines* lineares com seis nós foi superior aos demais modelos, devendo ser preferido para as avaliações genéticas da raça no Canadá.

Bignardi et al. (2008) estudaram funções B-*splines* para ajuste de dados de produção de leite no dia do controle de primeiras lactações de vacas da raça Holandês no Brasil. Os efeitos aleatórios genético-aditivo e de ambiente permanente foram modelados por meio de funções B-*splines* lineares, quadráticas ou cúbicas, com até sete nós. O resíduo foi considerado heterogêneo, contendo seis classes de variâncias. O modelo empregando uma função B-*spline* cúbica com o número de coeficientes de regressão aleatória igual a oito tanto para o efeito aleatório genético aditivo como de ambiente permanente foi o mais adequado para ajustar os dados. Para tal modelo, os valores de herdabilidade variaram de 0,20 e 0,40 ao longo da lactação.

Pereira et al. (2010) analisaram 38.268 registros de PLDC de 5.158 primeiras lactações de vacas da raça Gir Leiteiro. As PLDC foram agrupadas em dez classes mensais e analisadas por modelos de regressão aleatória, cujos efeitos aleatórios genético-aditivo e de ambiente permanente, foram

modelados utilizando-se funções *B-splines*. A modelagem da variância residual foi feita por meio de uma, quatro, seis ou dez classes. Os autores concluíram que o modelo considerando uma função *B-spline* quadrática, com cinco coeficientes de regressão aleatória para os efeitos aleatórios genético aditivo e de ambiente permanente e quatro classes de variâncias residuais é parcimonioso e ajusta-se aos registros de produção de leite no dia do controle de vacas Gir Leiteiro adequadamente.

Para a escolha do modelo que apresenta melhor ajuste para modelar a curva de produção de leite, comparam-se diferentes funções e consideram-se diferentes estruturas de variâncias residuais. Entretanto são necessários critérios para escolher o melhor modelo para a situação analisada. O modelo deve ser adequado para explicar o fenômeno, permitindo separar corretamente os efeitos em fixos e aleatórios genético aditivo e de ambiente permanente.

2.6. Heterogeneidade da variância residual

A adequada modelagem da variância residual (variância de ambiente temporário) é outro aspecto fundamental para uma melhor acurácia em avaliações genéticas sob modelos de regressão aleatória. De acordo com Araújo et al. (2010), a heterogeneidade de variância é uma forma primária da interação genótipo x ambiente. Sua existência pode comprometer o resultado final do processo de seleção, visto que se a fonte dessa heterogeneidade for devida a fatores ambientais, a seleção será realizada mais em função do ambiente do que propriamente do genótipo. Entretanto, de acordo com Silva (2011), os primeiros trabalhos sobre produção de leite

utilizando MRA desconsideravam a existência de uma estrutura de variância residual heterogênea ao longo da curva de lactação. Observou-se, contudo, que tal prática pode ocasionar equívocos nas estimativas de parâmetros genéticos. Isto pode ser ocasionado por diferenças de escala que influenciam alterações nas variâncias residuais ao longo da curva de lactação e/ou estar ligados a problemas de modelagem dos efeitos aleatórios.

A modelagem considerando as variâncias residuais heterogêneas para cada classe pode melhorar a partição da variação total, mas proporciona um aumento no número de parâmetros a serem estimados no processo de maximização da função de verossimilhança. O número elevado de parâmetros dos modelos, a estrutura de covariâncias e o fato de as matrizes de incidência serem mais densas que para os modelos convencionais podem dificultar a convergência na estimação dos componentes de variância, sendo entraves para a utilização dessa metodologia para avaliações genéticas em conjuntos com grandes volumes de dados (EL FARO e ALBUQUERQUE, 2003).

Para entender melhor suponha que se queira estimar a herdabilidade no início da lactação. Se a variância residual estimada for única para todo o período espera-se que a estimativa de herdabilidade esteja inflacionada, pois a variância fenotípica deste período é maior do que a variância fenotípica média. Espera-se também que a variância residual neste período seja superior à variância residual estimada considerando homogeneidade de variâncias. Este problema pode ser corrigido ao se utilizar uma estrutura de

variâncias residuais heterogêneas, adequadas às condições da análise, tais como característica, espécie, raça e clima (SILVA, 2011).

El Faro e Albuquerque (2003) utilizaram 86.598 registros de PLDC de 2.155 primeiras lactações de vacas da raça Caracu, truncadas aos 305 dias para comparar diferentes estruturas de variâncias residuais para modelos de regressão aleatória e constataram a necessidade de se considerar a heterogeneidade de variância residual.

Bignardi et al. (2008) analisaram 152.145 controles semanais de produção de leite de 7.317 primeiras lactações de vacas da raça Holandês. As PLDC's foram consideradas em 44 classes semanais de dias em lactação. O resíduo foi considerado heterogêneo, contendo 6 classes de variâncias, sendo que as semanas de lactação foram agrupadas da seguinte forma: 1, 2, 3, 4-6, 7-12, 13-44 semanas. De acordo com os autores, a desuniformidade no agrupamento das classes de variâncias residuais pode estar relacionada às diferenças nos níveis de produção ao longo da lactação, o que pode ser definido como efeito de escala.

Menezes et al. (2008) analisaram 10.238 registros semanais de PLDC provenientes de 388 primeiras lactações de cabras da raça Saanen para comparar MRA. Após definir o melhor modelo assumindo homogeneidade de variâncias residuais ao longo da lactação, contrastaram o uso de diferentes números de classes de variâncias residuais: homogeneidade, duas, três, quatro, cinco e seis classes ao longo da lactação. A utilização de seis classes ao longo da lactação foi indicada.

2.7. Avaliadores da qualidade do ajuste

Segundo Regazzi (2003) ao determinar a relação entre as variáveis, por meio de estudos de regressão, há o problema da definição da forma matemática que será ajustada para representar o fenômeno em estudo.

Frequentemente, ajusta-se mais de um modelo e com base nos resultados dos avaliadores de qualidade do ajuste conclui-se qual melhor modelo (OLIVEIRA, 2011).

Os critérios mais usados atualmente para contrastar MRA são: teste da razão de verossimilhança (TRV), 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 (SILVA, 2011). Isso pode ocorrer devido ao fato de modelos com maior número de parâmetros exigirem a utilização e inversão de matrizes de maiores dimensões e com maiores problemas de multicolinearidade.

Segundo Silveira (2010), a seleção do modelo mais adequado, com a utilização de vários avaliadores da qualidade do ajuste, fica mais precisa, uma vez que alguns dos avaliadores levam em consideração fatores como o grau de parametrização dos modelos comparados (baseia-se no princípio da parcimônia), e outros avaliadores consideram a análise da independência residual e outros.

O objetivo geral proposto neste trabalho foi estudar a variação da produção de leite no dia do controle de cabras das raças Alpina e Saanen por meio de modelos de regressão aleatória utilizando polinômios ortogonais de Legendre e polinômios segmentados do tipo B e compará-los.

Os objetivos específicos foram:

- a) Comparar os MRA ajustados para polinômios segmentados do tipo B e polinômios ortogonais de Legendre e identificar o melhor modelo para uma avaliação genética de caprinos;
- b) Avaliar diferentes estruturas de variâncias residuais para modelos de regressão aleatória, buscando minimizar o número de parâmetros a serem estimados e acelerar o processo de obtenção dos componentes de (co) variância.
- c) Estimar componentes de (co) variância e parâmetros genéticos e, prever valores genéticos para produção de leite ao longo da curva de lactação, usando MRA ajustados por polinômios ortogonais de Legendre e polinômios segmentados do tipo B;

3 – MATERIAL E MÉTODOS

3.1. Dados

O banco de dados utilizado foi proveniente do rebanho caprino do Departamento de Zootecnia da Universidade Federal de Viçosa (Viçosa, Minas Gerais). O rebanho é formado por animais das raças Alpina e Saanen. Todas as análises foram realizadas separadamente para as duas raças. Foram utilizados 17.356 registros de PLDC de 642 primeiras lactações de cabras da raça Alpina e 13.278 registros de 470 primeiras lactações de cabras da raça Saanen.

Os animais foram mantidos em baias coletivas sob o sistema de estabulação livre (*free stall*) e recebendo alimentação à base de silagem de milho e feno, além de mistura concentrada, fornecida conforme suas exigências nutricionais. Os controles leiteiros foram registrados semanalmente, oriundos de ordenha mecânica, realizada duas vezes ao dia. A PLDC representa a soma do leite coletado nestes dois períodos.

A organização, preparação e a análise de consistência do arquivo de dados foram realizadas utilizando-se o software SAS 9.1 (2003). Para melhor consistência dos dados foram mantidas no arquivo de análise, apenas primeiras lactações de cabras que tiveram o primeiro parto entre os 10 e 30 meses de idade; duração da lactação superior a 100 dias; partos simples, duplos ou triplos, dias em lactação superior a cinco, e intervalos de controles leiteiros inferiores a 28 dias. As lactações foram truncadas aos 270 dias.

Para a recodificação do arquivo de genealogia, foi utilizado o software RENPED (SILVA, 2011). A matriz dos numeradores dos coeficientes de

parentesco utilizada para as análises dos animais das duas raças continha 4.767 indivíduos.

3.2. Análises

3.2.1. Efeitos fixos

Após a edição dos dados, avaliou-se a influência dos efeitos que poderiam influenciar a PLDC. Os efeitos testados foram: ano e estação de parto (de 2000 a 2011, para a raça Alpina e 2001 a 2011 para a raça Saanen e estação 1, de março a setembro e 2 para os demais meses), agrupamento genético, tipo de parto (até três filhotes) e idade da cabra ao parto como covariável, considerando-se os efeitos linear e quadrático.

As análises foram realizadas separadamente para as raças Alpina e Saanen. No arquivo da raça Alpina, apenas foram utilizados animais com composição genética desta raça igual ou superior a 50% e as classes genéticas foram agrupadas da seguinte forma: 1 = 50% raça Alpina; 2 = superior a 50% e igual ou inferior a 75%; 3 = superior a 75% e igual ou inferior a 93% e 4 = superior a 93%. As classes de agrupamentos genéticos para a raça Saanen foram formadas da mesma forma.

O modelo utilizado para testar a influência dos fatores sobre a PLDC foi:

$$y_{ijkl} = \mu + T_i + G_j + AE_k + b_1(P_{ijkl} - \bar{P}) + b_2(P_{ijkl} - \bar{P})^2 + \varepsilon_{ijkl}$$

em que:

y_{ijkl} = valor observado da PLDC na cabra l, tipo de parto i, agrupamento genético j, ano-estação k;

μ = média populacional da PLDC;

T_i = efeito fixo do tipo de parto i ;

G_j = efeito fixo do agrupamento genético j ;

AE_k = efeito fixo do ano-estação k ;

b_1 e b_2 = coeficientes linear e quadrático de regressão da característica analisada em função da idade da cabra ao primeiro parto (ICPP);

P_{ijkl} = idade da cabra l ao primeiro parto;

\bar{P} = média amostral da ICPP;

ε_{ijkl} = erro aleatório;

As análises estatísticas foram realizadas empregando-se o procedimento *Generalized Linear Model* (Proc GLM) do *software Statistics Analysis System* (SAS, 2003).

3.2.2. Polinômios Ortogonais de Legendre

As análises foram realizadas por meio de modelos unicaracterísticos de regressão aleatória. Para todos os modelos foram incluídos os efeitos aleatórios genético aditivo e de ambiente permanente de animal e o resíduo.

O modelo animal unicaracterístico de regressão aleatória, é descrito de forma generalizada como:

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

em que: y_{ij} corresponde à produção de leite no controle j da cabra i ; EF_i consiste no conjunto de efeitos fixos, constituído por agrupamento genético, 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 ; k_b , k_α , k_γ 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 a cada observação não explicados pelos fatores de correção nem pelas regressões do modelo.

Os modelos testados continham combinações distintas de ordens de ajuste para os polinômios das curvas, média de lactação da população (fixa), genética aditiva direta (aleatória) e de ambiente permanente (aleatória), além de número diferente de classes de variâncias residuais. Os primeiros modelos avaliados variaram da ordem três até a sexta para as curvas fixa, genética aditiva direta e ambiente permanente, considerando homogeneidade de variâncias. Apenas foram consideradas funções de mesma ordem de ajuste para permitir possibilidades iguais de variações de ambos os componentes, conforme recomendação de López-Romero et al. (2004).

De posse do modelo mais adequado considerando homogeneidade de variância residual, avaliou-se duas formas de estruturas de variâncias residuais: variância residual constante ao longo da lactação ou variância residual constante dentro, mas heterogênea entre classes de dias em lactação. Quando heterogêneas, foram consideradas as seguintes classes: duas (do 6º ao 160º e do 161º ao 270º dia de lactação), três (do 6º ao 60º, do 61º ao 150º e do 151º ao 270º dia de lactação), quatro (do 6º ao 50º, 51º

ao 100^o, 101^o a 200^o e 201^o a 270 dias de lactação), cinco (do 6^o ao 30^o, 31^o ao 90^o, 91^o ao 141^o, 141^o ao 200^o e 201^o ao 270^o dia de lactação) e seis (do 6^o ao 30^o, 31^o ao 90^o, 91^o a 120^o, 121^o a 160^o, 161^o a 210^o e 211^o a 270^o dias de lactação), totalizando seis diferentes modelos a serem avaliados nessa segunda etapa. As classes de variâncias residuais foram formadas com base em uma inspeção gráfica dos registros de PLDC. É interessante frisar que as classes de variâncias residuais não precisam ser separadas em intervalos iguais.

A notação dos modelos ajustados pelos polinômios ortogonais de Legendre segue o padrão: LEGkHx, com k referindo-se à ordem do polinômio para modelar a curva média da população e os efeitos genético aditivo e de ambiente permanente e x é o número de classes de variâncias residuais. Por exemplo, um modelo com curva de regressão média, genética aditiva e de ambiente permanente de ordem 3 e que considerou a variância residual heterogênea contendo quatro classes, será identificado como LEG3H4.

3.2.3. – Polinômios Segmentados do Tipo B (B-splines)

As análises foram realizadas por meio de modelos unicaracterísticos de regressão aleatória. Para todos os modelos foram incluídos os efeitos aleatórios, genético aditivo e de ambiente permanente de animal. Foram considerados como efeitos fixos, o grupo contemporâneo (ano–estação de parição), agrupamento genético, tipo de parto (simples, duplo ou triplo) e os efeitos linear e quadrático da covariável idade da cabra ao parto.

A forma matricial do modelo animal, unicaracterístico, de regressão aleatória é dada por:

$$y = Xb + Za + Wc + e,$$

em que: y é o vetor das N observações, medidas em N_d animais; b é o vetor de efeitos sistemáticos e dos coeficientes da regressão fixa; a é o vetor de coeficientes de regressão aleatória genético aditivo; c é o vetor dos coeficientes de regressão aleatória de ambiente permanente; e é o vetor de efeito aleatório residual; e X , Z e W são as matrizes de incidência dos efeitos fixos e aleatórios genético aditivo e de ambiente permanente, respectivamente.

As pressuposições em relação aos componentes do modelo são:

$$E \begin{bmatrix} y \\ a \\ c \\ e \end{bmatrix} = \begin{bmatrix} Xb \\ 0 \\ 0 \\ 0 \end{bmatrix} e,$$

$$V(a) = K_A \otimes A;$$

$$V(c) = K_{AP} \otimes I_{N_d} e,$$

$$V(e) = R, \text{ em que:}$$

K_A e K_C são as matrizes de (co) variâncias entre os coeficientes de regressão aleatórios genético-aditivos e de ambiente permanente, respectivamente;

A é a matriz dos numeradores dos coeficientes de parentesco entre os indivíduos;

I_{N_d} é uma matriz identidade;

R representa uma matriz contendo variâncias residuais.

Os efeitos aleatórios, genético aditivo e de ambiente permanente e a curva média de lactação foram modelados por meio de funções *B-splines* (linear, quadrático e cúbico), com nós equidistantes.

As *B-splines* de grau $p=0$ têm valores iguais à unidade para todos os pontos (t) em um dado intervalo, e zero caso contrário. Para o k -ésimo segmento dado por nós T_k e T_{k+1} com $T_k \leq T_{k+1}$:

$$B_{k,0}(t) = \begin{cases} 1, & \text{se } T_k \leq t < T_{k+1}; \\ 0, & \text{caso contrário.} \end{cases}$$

Já as funções de grau $B_{k,p}$, para $p>0$, podem ser representadas por:

$$B_{k,p}(t) = \frac{t - T_k}{T_{k+p} - T_k} \cdot B_{k,p-1}(t) + \frac{T_{k+p+1} - t}{T_{k+p+1} - T_{k+1}} \cdot B_{k+1,p-1}(t)$$

Os segmentos individuais foram considerados por meio de polinômios de ordem linear (L), quadrático (Q) e cúbico (C), com funções bases de grau $p = 1, 2$ e 3 , unidos em pontos determinados nós, o que permite que a função se torne contínua. Polinômios de mesmo grau foram considerados no modelo para a curva fixa e para os efeitos aleatórios genético aditivo e de ambiente permanente. Primeiramente foram avaliadas as funções *B-splines* para todas as curvas, considerando homogeneidade de variâncias residuais. De posse do modelo mais adequado sob homogeneidade de variâncias residuais, avaliou-se duas formas de estruturas de variâncias residuais: variância residual constante ao longo da lactação ou variância residual constante dentro, mas heterogênea entre classes de dias em lactação. Quando heterogênea, foram consideradas as seguintes classes: duas (do 6º ao 160º e do 161º ao 270º dia de lactação), três (do 6º ao 60º, do 61º ao

150° e do 151° ao 270° dia de lactação), quatro (do 6° ao 50°, 51° ao 100°, 101° a 200° e 201° a 270 dias de lactação), cinco (do 6° ao 30°, 31° ao 90°, 91° ao 141°, 141° ao 200° e 201° ao 270° dia de lactação) e seis (do 6° ao 30°, 31° ao 90°, 91° a 120°, 121° a 160°, 161° a 210° e 211° a 270° dias de lactação).

Os m nós foram escolhidos para dividir as semanas de lactação em $m-1$ intervalos equidistantes e os nós externos foram localizados no 6° e 270° dia de lactação para todos os modelos. O mesmo número de nós foi ajustado tanto para o efeito genético aditivo quanto para o ambiente permanente e para a curva fixa da população. Os números de coeficientes de regressão aleatória para modelar a trajetória das funções bases linear, quadrática e cúbica, são dados, respectivamente, por m , $m+1$ e $m+2$. Os nós extremos não interferem no número de parâmetros do modelo.

A citação dos MRA seguirá o padrão: “BSXkHy”, em que X= L, Q e C são os graus do segmento polinomial linear, quadrático e cúbico, respectivamente, k corresponde ao número de nós para a curva média da população e para os efeitos genético aditivo e de ambiente permanente e y, ao número de classes de variâncias residuais. Portanto, um modelo tendo a seguinte representação BSL3H3: constará de uma função B-spline linear, considerando três classes de variâncias residuais e três nós tanto para a curva média como para os efeitos aleatórios genético aditivo e de ambiente permanente, com dois segmentos ($k - 1$) e três nós ($k = 3$), uma vez que existem dois nós em pontos extremos, no 6° e 270° dia de lactação após o parto.

Os componentes de variância foram estimados pelo Método de Máxima Verossimilhança Restrita (REML), utilizando-se o programa Wombat (MEYER, 2006).

Os resultados obtidos pelos modelos de regressão aleatória, ajustados por funções *B-splines*, foram também comparados com o melhor modelo obtido utilizando polinômios ortogonais de Legendre.

3.2.4. Comparação dos modelos

Foram avaliados polinômios ortogonais de Legendre e polinômios segmentados do tipo B de diferentes ordens e estruturas de variâncias residuais. Para a escolha dos que se ajustaram melhor aos dados, os modelos foram comparados usando-se o logaritmo da função de máxima verossimilhança restrita (Log L) e as formas de máxima verossimilhança restrita dos Critérios de Informação de Akaike (AIC) (AKAIKE, 1973) e Bayesiano de Schwarz (BIC) (SCHWARZ, 1978).

Os critérios AIC e BIC permitem uma comparação entre modelos e penalizam aqueles com maior número de parâmetros, sendo que o BIC atribui uma penalidade mais rigorosa. Menores valores para AIC e BIC e, maiores para Log L, indicam melhor ajuste do modelo. Os critérios de informação podem ser representados como:

$$AIC = - 2 \text{ Log } L + 2 p,$$

$$BIC = - 2 \text{ Log } L + p \text{ Log } (n - r),$$

em que, p é o número de parâmetros do modelo, n é o número total de informações, r é o posto da matriz de incidência de efeitos fixos no modelo e $\text{Log } L$ é o logaritmo da função de máxima verossimilhança restrita.

Para comparação entre modelos completos e reduzidos também foi utilizado o Teste da Razão de Verossimilhança (TRV) (WOLFINGER, 2003), ao nível de 1% de probabilidade. O TRV compara as diferenças entre as funções de máxima verossimilhança linearizadas de cada modelo, com o valor obtido de uma distribuição qui-quadrado com o número de graus de liberdade igual à diferença no número de parâmetros entre os modelos.

Os componentes de (co) variância e parâmetros genéticos estimados ao longo da curva de lactação obtidos pelos modelos que proporcionaram melhores ajustes também foram inspecionados e comparados.

Além disso, foi feito um acompanhamento visando verificar a ocorrência de problemas de convergência fora do espaço paramétrico ou que gerasse estimativas não condizentes com a biologia da característica analisada.

4 – RESULTADOS E DISCUSSÃO

4.1. Análise de efeitos fixos (raças Alpina e Saanen)

As análises foram realizadas separadamente para as duas raças. Todos os efeitos testados (agrupamento genético, ano-estação de parto, idade da cabra ao parto, efeito linear e quadrático, e tipo de parto) foram significativos, para as duas raças, ao nível de 1% de significância e, portanto foram incluídos no modelo para estudo da produção de leite no dia do controle (PLDC).

A influência do efeito ano-estação já era esperada, pois as flutuações climáticas, mais especificamente a variação de pluviosidade, temperatura e umidade do ar, entre e dentro dos anos estudados causam variações no desempenho por meio da estacionalidade na disponibilidade de alimentos, em termos qualitativos e quantitativos, mesmo com o manejo alimentar adotado no setor experimental e, além disso, exercem influência sobre o conforto dos animais em confinamento.

Ao longo do processo de seleção dos animais para determinadas características pode ocorrer uma diferenciação genética entre raças que estejam sendo selecionadas para uma mesma função, como produção de leite. Pode ocorrer a fixação de determinadas formas alélicas em uma raça e em outra raça serem fixados alelos distintos, já que elas podem ter passado por processos diferenciados de seleção natural e artificial. Sendo assim, o efeito significativo de agrupamento genético pode ser atribuído à constituição genética do rebanho estudado, pois nas análises não foram considerados

apenas animais geneticamente puros e como há diferenças produtivas entre as raças, há necessidade de se corrigir para esse efeito.

O efeito de tipo de parto influenciou a PLDC possivelmente devido ao fato de neste trabalho ter sido considerado apenas as primeiras lactações e durante a fase gestacional as cabritas ainda estarem em fase de crescimento. O dispêndio de nutrientes para a gestação de um ou mais filhotes pode afetar o desenvolvimento da cabra, principalmente suas reservas energéticas e com isso influenciar na produção de leite subsequente, como ocorrido no presente estudo.

A influência da idade da cabra ao parto pode ser explicada pelo fato de que cabras que parem mais velhas já estabilizaram o crescimento e com isso não há competição entre produção de leite e desenvolvimento corporal, enquanto que esse efeito é verificado nas que se reproduzem mais jovens, pois ainda não atingiram a maturidade fisiológica.

4.2. Raça Alpina

Na Tabela 1 são apresentados o número de observações, média e desvio padrão para a característica PLDC. Os valores apresentados se encontram dentro dos padrões aceitáveis para caprinos de aptidão leiteira. Soares Filho et al. (2001), encontraram média de 2,01 Kg/dia para cabras da raça Alpina de rebanhos do Distrito Federal e Brito et al. (2009) encontraram média de 1,76 Kg/dia. Embora os valores observados de produção de leite sejam razoáveis para rebanhos brasileiros, ainda há muito a se melhorar para atingir os patamares produtivos de países como a França que vêm investindo na seleção genética da espécie há muitos anos.

Tabela 1 – Número de registros (NR), média e desvio-padrão (DP) para cada produção de leite na semana do controle da primeira lactação de cabras da raça Alpina.

Semana	NR	Média (Kg)	DP	Semana	NR	Média (Kg)	DP
2	448	2,16	0,85	21	479	2,16	0,93
3	527	2,32	0,91	22	464	2,07	0,92
4	561	2,39	0,96	23	456	2,09	0,91
5	558	2,52	1,00	24	437	1,98	0,90
6	565	2,55	1,01	25	440	1,98	0,88
7	542	2,61	1,05	26	432	1,91	0,88
8	563	2,60	1,04	27	417	1,85	0,85
9	563	2,63	1,05	28	415	1,81	0,84
10	563	2,66	1,05	29	390	1,80	0,82
11	542	2,59	1,04	30	393	1,77	0,79
12	533	2,52	1,02	31	382	1,75	0,80
13	511	2,50	0,97	32	372	1,69	0,80
14	520	2,51	0,99	33	364	1,67	0,80
15	518	2,48	1,01	34	337	1,70	0,75
16	524	2,45	0,99	35	340	1,67	0,81
17	526	2,37	1,00	36	316	1,64	0,83
18	513	2,30	0,97	37	317	1,65	0,82
19	521	2,27	0,98	38	294	1,65	0,83
20	480	2,23	0,92	39	200	1,67	0,83

Na Figura 1 é apresentada a PLDC em função da semana de lactação. A produção de leite variou conforme uma curva padrão de lactação para raças especializadas, iniciando com 2,16 Kg no primeiro controle leiteiro, aumentando até 2,66 Kg na décima semana após o parto e reduzindo gradualmente nos controles leiteiros subsequentes até 1,67 na 39ª semana de lactação.

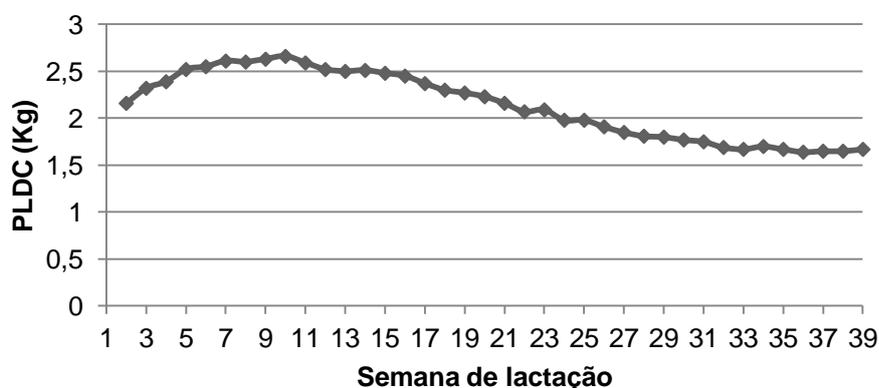


Figura 1 – Produção de leite ao longo da lactação de cabras da raça Alpina.

4.2.1. Ajuste e comparação de modelos

a) Polinômios Ortogonais de Legendre

O emprego de modelos de regressão aleatória por meio de funções polinomiais de Legendre requer, inicialmente, a definição da ordem mais apropriada para a regressão média da população e para os efeitos aleatórios considerados no modelo de análise (genético aditivo e de ambiente permanente) e a avaliação da necessidade de considerar a heterogeneidade de variância residual.

Na Tabela 2 estão apresentados os modelos de regressão aleatória ajustados por polinômios ortogonais de Legendre considerando homogeneidade de variâncias residuais, além dos critérios utilizados para contrastá-los.

Tabela 2 – Modelos, número de parâmetros (p), AIC, BIC, Log L e TRV para os polinômios ortogonais de Legendre, considerando homogeneidade de variâncias residuais.

Modelo*	p	Log L	AIC	BIC	TRV
LEG3H1	13	2374,138	-4722,276	-4621,400	**
LEG4H1	21	2714,039	-5386,076	-5223,122	**
LEG5H1	31	2941,077	-5820,154	-5579,608	**
LEG6H1	43	3050,944	-6015,888	-5682,228	completo

** = significativo a 1%; *Modelo: LEGxHy, em que x corresponde à ordem dos polinômios para a curva fixa e para os efeitos aleatórios genético aditivo e de ambiente permanente e y, ao número de classes de variâncias residuais; valores em negrito indicam o melhor modelo com base no AIC e BIC.

De acordo com os resultados apresentados na Tabela 2, verificou-se que os valores de Log L, AIC e BIC melhoraram com o aumento na ordem dos polinômios e de acordo com o TRV foram encontradas diferenças significativas ($P < 0,01$) entre o modelo completo e todos os modelos reduzidos.

De forma geral, maiores valores para o Log L sempre serão esperados para os modelos mais parametrizados, pois estes são mais flexíveis, se ajustando melhor aos dados. Entretanto, devem ser utilizados outros critérios que permitam a escolha de modelos mais parcimoniosos como os critérios AIC e BIC, que vêm se tornando padrão para comparação de modelos.

Dentre os modelos que consideraram homogeneidade de variâncias residuais, o modelo LEG6H1 foi o que melhor se ajustou aos dados segundo os critérios adotados para compará-los. Apesar disso, é importante a avaliação da necessidade de considerar a heterogeneidade de variâncias residuais ao longo da lactação, pois pode haver melhorias na partição da variação total e com isso melhores estimativas dos parâmetros genéticos.

Com o aumento no número de classes há mais parâmetros a serem estimados no processo de maximização da função de verossimilhança e modelos mais parametrizados podem possuir maiores problemas de convergência e de estimação, o que pode dificultar a utilização dos modelos de regressão aleatória em avaliações genéticas com grande volume de dados. Neste sentido, busca-se a utilização de modelos mais parcimoniosos, sem perda da qualidade de ajuste dos efeitos aleatórios considerados. Os resultados dos modelos considerando heterogeneidade de variâncias residuais estão apresentados na Tabela 3.

Verificou-se que o modelo que considera homogeneidade de variâncias residuais obteve o pior ajuste dentre todos os modelos avaliados (Tabela 3). Este resultado está de acordo com os encontrados por Breda et al. (2006), Assis et al. (2006), Menezes (2008) e Silva (2011).

Tabela 3 – Modelos, número de parâmetros (p), AIC, BIC, Log L e TRV para os polinômios ortogonais de Legendre, considerando heterogeneidade de variâncias residuais.

Modelo*	p	Log L	AIC	BIC	TRV
LEG6H1	43	3050,944	-6015,888	-5682,228	**
LEG6H2	44	3142,035	-6196,070	-5854,652	**
LEG6H3	45	3155,987	-6221,974	-5872,796	**
LEG6H4	46	3153,432	-6214,864	-5857,928	**
LEG6H5	47	3179,123	-6264,246	-5899,548	ns
LEG6H6	48	3182,000	-6268,000	-5895,544	completo

** = significativo a 1%; ns = não-significativo; *Modelo: LEGxHy, em que x corresponde à ordem dos polinômios para a curva fixa e para os efeitos aleatórios genético aditivo e de ambiente permanente e y, ao número de classes de variâncias residuais; valores em negrito indicam o modelo escolhido.

Assumir homogeneidade de variâncias residuais ao longo da curva de lactação pode criar distorções consideráveis na partição da variância total, uma vez que a variância residual é influenciada por muitos fatores que podem diferir ao longo da lactação, principalmente o efeito de escala da produção que é maior no estágio inicial da lactação do que no final. Isto pode ser verificado através do resultado da divisão em apenas duas classes de variâncias residuais, o qual já proporcionou uma grande redução nos valores de AIC e BIC, indicando melhores ajustes. Além disso, o aumento no número de classes melhorou de forma considerável a modelagem dos efeitos aleatórios envolvidos na análise. Estes resultados estão de acordo com os apresentados por Breda et al. (2006) e Silva (2011) em caprinos da raça Alpina, que também encontraram essa tendência de melhoria de ajuste com o aumento do número de classes do resíduo.

O modelo LEG6H6, apresentou melhor ajuste com base nos critérios AIC e Log L. Entretanto, apresentou um valor de BIC um pouco inferior ao modelo LEG6H5. Isto pode ser explicado devido ao fato de o critério BIC aplicar maior penalização aos modelos mais parametrizados e o modelo LEG6H6 possui um parâmetro a mais do que o modelo anterior. Além disso,

não foram encontradas diferenças significativas ao nível de 1% de probabilidade, com base no TRV entre esses dois modelos. Como um maior número de parâmetros gera maior demanda computacional, podendo ocasionar maiores dificuldades de convergência dos algoritmos, o modelo que considera cinco classes de variâncias residuais (LEG6H5) é o mais indicado em relação a todos os outros avaliados anteriormente.

Sarmiento et al. (2006) também trabalhando com dados de cabras da raça Alpina deste mesmo rebanho, coletados entre 1997 e 2004, indicaram o modelo de quinta ordem para o efeito genético aditivo e de sétima ordem para o efeito de ambiente permanente como o melhor. Entretanto, sugerem que a utilização de um modelo menos parametrizado, de quarta ordem para o efeito genético aditivo e de sexta ordem para o efeito de ambiente permanente, apresentaria resultados semelhantes.

Silva (2011) analisou 20.710 registros de PLDC de 667 cabras da raça Alpina com o objetivo de identificar o melhor MRA utilizando polinômios ortogonais de Legendre. Os modelos testados continham combinações de distintas ordens de ajustes para os polinômios das curvas fixa (2 a 5), genética aleatória (1 a 7) e ambiente permanente (1 a 7) e número de classes de variâncias residuais (2, 4, 5 e 6). O MRA utilizando polinômios ortogonais de Legendre mais indicado para avaliação genética da PLDC de caprinos da raça Alpina foi o que considera a curva fixa de ordem 4, a curva de efeitos genéticos aditivos de ordem 2, a curva de efeitos de ambiente permanente de ordem 7 e ao menos cinco classes de variâncias residuais.

Após a escolha do melhor modelo utilizando polinômios ortogonais de Legendre, foram inspecionados os componentes de variância estimados

peelo modelo e as estimativas de herdabilidade para todos os pontos da lactação e as correlações genéticas entre as produções de leite nos dias 6, 71, 137, 203 e 270.

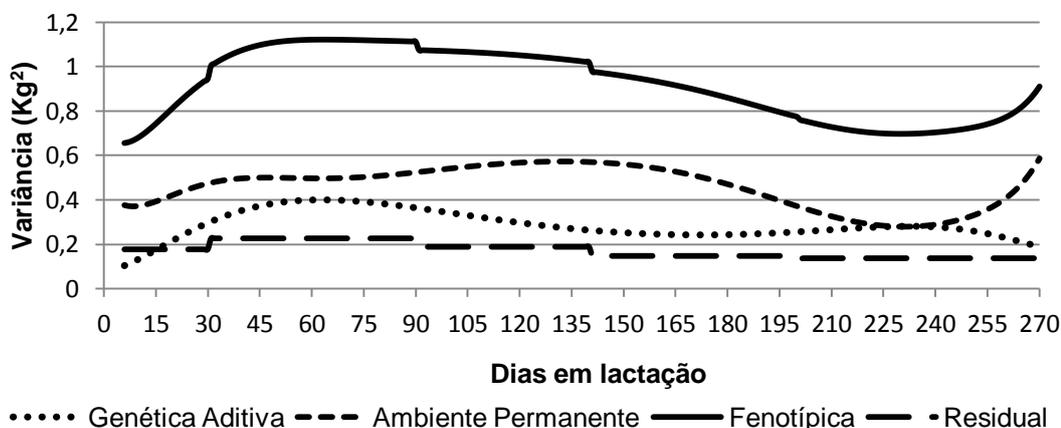


Figura 2 – Estimativas de variâncias obtidas pelo modelo LEG6H5.

Os maiores valores de variância genética aditiva e residual foram observados na fase do pico de produção, que ocorre por volta dos 40 dias após o parto (Figura 2). Uma das possíveis explicações seria o efeito de escala, já que nessa fase são observadas as maiores produções da lactação. Esses resultados diferem dos encontrados por Silva (2011), em trabalho com caprinos da raça Alpina, que observou uma variância genética crescente ao longo da lactação.

A variância fenotípica e de ambiente permanente parecem estar subestimadas no início da lactação e superestimadas no final. Isso pode ter ocorrido devido ao menor número de dados no início e final da lactação, o que não representa grandes problemas para a predição dos valores genéticos, mas cria uma extrapolação na predição de valores associados aos efeitos de ambiente permanente dos animais que perderam dados no início ou final da lactação. Conforme pode ser observado na Figura 2, as estimativas de variâncias no período final parecem piores do que as

estimativas do início da lactação, e é justamente no final da lactação que houve uma redução mais acentuada no número de registros. Menezes et al. (2010a) também encontraram maiores estimativas de ambiente permanente no final da lactação.

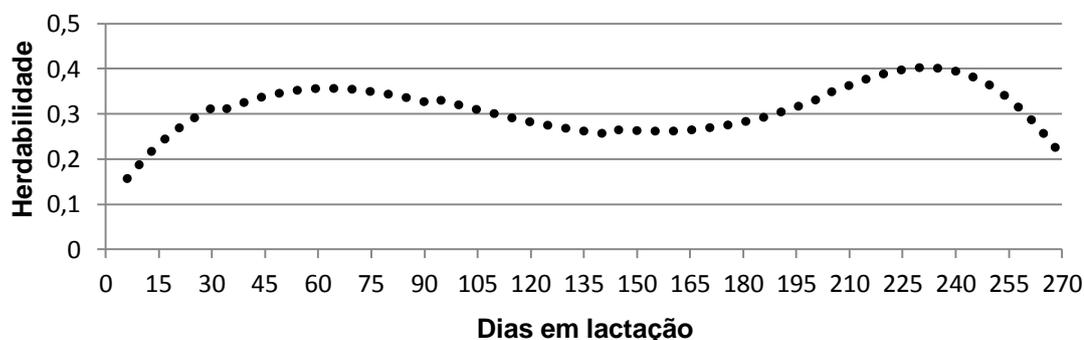


Figura 3 – Estimativas de herdabilidade obtidas pelo modelo LEG6H5.

As estimativas de herdabilidade variaram de 0,15 a 0,40 ao longo da lactação (Figura 3). Verifica-se que as estimativas de herdabilidade foram menores no início e final da lactação. Isto possivelmente ocorreu devido à maior influência dos efeitos ambientais nessas fases. No início da lactação, as cabras ainda não se adaptaram ao processo de ordenha, sofrem um estresse pós-parto devido ao balanço energético negativo e no final da lactação, muitas delas já estão em estágio avançado da gestação. Além disso, conforme já mencionado, nesses dois extremos há um menor número de dados, o que também pode ocasionar erros nas estimativas das variâncias e conseqüentemente nas estimativas de herdabilidade e correlações genéticas, principalmente porque a perda de dados nestes períodos leva a uma extrapolação nas estimativas de ambiente permanente para alguns indivíduos. Resultados semelhantes foram relatados por Sarmiento et al. (2008) que encontraram estimativas mais altas no meio da

lactação. Entretanto, Silva (2011) encontrou maiores estimativas no terço final e Menezes et al. (2010a), encontraram maiores valores no início da curva de lactação.

Na Tabela 4 são apresentadas as estimativas de correlações genéticas, obtidas pelo modelo LEG6H5, entre as produções de leite nos dias 6, 71, 137, 203 e 270. As correlações genéticas nesse estudo, obtidas pelo modelo LEG6H5, entre os dias avaliados, variaram de 0,21 a 0,95. Silva (2011) encontrou estimativas de correlações genéticas variando de 0,98 a 1,00. As correlações genéticas foram menores à medida que os intervalos entre as PLDC's aumentaram. Este resultado está de acordo com Araújo et al. (2006), que em um trabalho com bovinos da raça Holandês também encontraram estimativas de correlações genéticas altas e maiores à medida que os controles eram mais próximos no tempo. Uma possível explicação para isso seria que tanto o grupo gênico quanto a expressão dos genes ligados à produção de leite varia ao longo da lactação, sendo esta expressão mais semelhante para controles leiteiros mais adjacentes.

Tabela 4 – Estimativas de correlações genéticas para PLDC's obtidas pelo modelo LEG6H5.

PLDC	6	71	137	203	270
6		0,95	0,84	0,71	0,25
71			0,94	0,75	0,21
137				0,88	0,40
203					0,78
270					

Apesar das correlações genéticas entre produções mais distantes terem sido de menores magnitudes, todas as estimativas foram positivas indicando que a seleção dos animais para a produção de leite em um dia

qualquer da lactação trará respostas positivas para todos os demais pontos da curva.

b) Polinômios segmentados do tipo B (*B-splines*)

Após o estudo dos modelos utilizando polinômios ortogonais de Legendre, avaliou-se o ajuste de modelos de regressão aleatória utilizando polinômios segmentados do tipo B. Na Tabela 5 são apresentados os modelos que foram testados considerando homogeneidade de variâncias residuais.

Tabela 5 – Modelos, número de parâmetros (p), AIC, BIC, Log L e TRV para os polinômios segmentados do tipo B, considerando homogeneidade de variâncias residuais.

Modelo*	p	Log L	AIC	BIC	TRV
Linear					
BSL3H1	13	2201,291	-4376,582	-4275,706	**
BSL4H1	21	2631,513	-5221,026	-5058,074	**
BSL5H1	31	2876,156	-5690,312	-5449,766	**
BSL6H1	43	3066,300	-6046,600	-5712,942	completo
Quadrático					
BSQ3H1	13	2339,923	-4653,846	-4552,978	**
BSQ4H1	21	2711,716	-5381,432	-5218,480	**
BSQ5H1	31	2952,442	-5842,884	-5610,338	**
BSQ6H1	43	3092,906	-6099,812	-5766,154	completo
Cúbico					
BSC4H1	21	2717,819	-5393,638	-5230,686	**
BSC5H1	31	2958,138	-5852,276	-5611,728	**
BSC6H1	43	3079,741	-6073,482	-5739,824	completo

** = significativo a 1%; *Modelo: BSXkHy, em que X = L, Q e C são os graus do segmento polinomial linear, quadrático e cúbico, respectivamente, k corresponde ao número de nós para a curva média da população e para os efeitos genético aditivo e de ambiente permanente e y, ao número de classes de variâncias residuais; valores em negrito indicam o melhor modelo com base no AIC e BIC.

O software WOMBAT (MEYER, 2006) não permite a avaliação de modelos BSC3 (três nós), pois para funções *B-splines* do tipo cúbico, espera-se que ao menos um parâmetro influencie a variável dependente em

pelo menos três segmentos. Para tal, seriam necessários ao menos quatro nós.

De acordo com os resultados apresentados na Tabela 5, independentemente da função *B-spline* empregada (linear, quadrática ou cúbica), quando se aumentou o número de nós, tanto o AIC quanto o BIC apresentaram valores absolutos superiores, indicando melhores ajustes. Considerando apenas o número de nós, os valores de AIC e BIC foram sempre maiores, em módulos, para polinômios de segundo grau (quadrático), indicando que estes se ajustaram melhor aos dados.

Dentre os polinômios segmentados do tipo B considerando homogeneidade de variâncias residuais, o modelo que melhor se ajustou aos dados foi o BSQ6H1. No entanto, as diferenças obtidas entre os critérios dos modelos BSL6H1, BSQ6H1 e BSC6H1 foram relativamente pequenas e o TRV não permite comparação entre os mesmos, por não haver graus de liberdade suficientes, pois o número de parâmetros nos três modelos é o mesmo (43). Sendo assim, também foram avaliados números diferentes de classes de variâncias residuais para os polinômios segmentados do tipo B de graus 1, 2 e 3 (BSL6, BSQ6 e BSC6).

Dentre os modelos com o mesmo número de parâmetros, mas com diferentes graus de função *B-spline* (linear, quadrático e cúbico), aqueles que utilizaram a função *B-spline* quadrática mostraram-se sempre superiores de acordo com os critérios adotados (Tabela 6).

Além disso, para todas as ordens avaliadas, os modelos que apresentaram melhores ajustes foram os que consideraram cinco classes de variâncias residuais, assim como o melhor modelo ajustado por polinômios

ortogonais de Legendre (LEG6H5), que também considerou cinco classes. Esses resultados indicam que esse é realmente o número de classes mais adequado a ser utilizado nas análises de PLDC para caprinos leiteiros da raça Alpina. É importante frisar que a divisão mais adequada das classes de variâncias residuais também influenciará no número de classes a serem utilizadas.

Tabela 6 – Modelos, número de parâmetros (p), AIC, BIC, Log L e TRV para os polinômios segmentados do tipo B, considerando heterogeneidade de variâncias residuais.

Modelo	p	Log L	AIC	BIC	TRV
Linear					
BSL6H1	43	3066,300	-6046,600	-5712,942	**
BSL6H2	44	3155,888	-6223,776	-5882,358	**
BSL6H3	45	3192,276	-6294,552	-5945,374	**
BSL6H4	46	3184,473	-6276,946	-5920,010	**
BSL6H5	47	3209,369	-6324,738	-5960,042	ns
BSL6H6	48	3208,161	-6320,322	-5947,866	completo
Quadrático					
BSQ6H1	43	3092,906	-6099,812	-5766,154	**
BSQ6H2	44	3185,676	-6283,352	-5941,934	**
BSQ6H3	45	3202,330	-6314,660	-5965,482	**
BSQ6H4	46	3197,208	-6302,416	-5945,478	**
BSQ6H5	47	3223,592	-6353,184	-5988,488	ns
BSQ6H6	48	3224,676	-6353,352	-5980,896	completo
Cúbico					
BSC6H1	43	3079,741	-6073,482	-5739,824	**
BSC6H2	44	3171,470	-6254,940	-5913,522	**
BSC6H3	45	3185,888	-6281,776	-5932,598	**
BSC6H4	46	3182,311	-6272,622	-5915,684	**
BSC6H5	47	3208,625	-6323,250	-5958,554	ns
BSC6H6	48	3210,447	-6324,894	-5952,438	completo

** = significativo a 1%; ns = não-significativo; *Modelo: BSXkHy, em que X = L, Q e C são os graus do segmento polinomial linear, quadrático e cúbico, respectivamente, k corresponde ao número de nós para a curva média da população e para os efeitos genético aditivo e de ambiente permanente e y, ao número de classes de variâncias residuais; valores em negrito indicam o melhor modelo com base no AIC e BIC.

Dentre todos os modelos avaliados utilizando polinômios segmentados do tipo B, o BSQ6H5 foi o que melhor se ajustou aos dados de acordo com os critérios adotados. O fato de o modelo BSQ6H5 ter se ajustado melhor aos dados que os modelos BSL6H5 e BSC6H5 pode ter ocorrido devido à forma como cada um deles considera os pontos que influenciam os parâmetros de cada nó. Os modelos utilizando *B-splines* lineares praticamente desconsideram na estimação de um ponto as influências de pontos mais distantes, enquanto que as funções *B-splines* cúbicas são as que mais levam em consideração a influência entre pontos mais distantes. As funções *B-splines* lineares utilizam um menor número de pontos (registros) para a estimação dos coeficientes. Entretanto, apresentam menores problemas de multicolinearidade. No caso das funções *B-splines* cúbicas mais pontos são utilizados para a estimação dos coeficientes. Contudo, além da maior multicolinearidade, registros do final da lactação podem influenciar pontos do início e vice-versa. Isto não é desejável, pois os genes que influenciam o pico de produção, a duração da lactação e a persistência podem não serem os mesmos. Nesse estudo, para os dados de PLDC de animais da raça Alpina, as funções *B-splines* quadráticas equilibraram melhor as vantagens e desvantagens de cada ordem.

Os resultados das Tabelas 3 e 6 mostram que os modelos que consideraram variâncias residuais homogêneas foram os que apresentaram piores ajustes aos dados de acordo com os critérios utilizados. Isto evidencia a necessidade de considerar uma estrutura heterogênea de variâncias residuais, uma vez que as variâncias tendem a diferir ao longo da lactação. Resultados semelhantes foram obtidos por Bignardi et al. (2011), que

avaliaram modelos para ajustes de curvas de lactação de vacas da raça Holandês e segundo os autores, de acordo com os critérios AIC e BIC, os modelos que continham variâncias residuais homogêneas proporcionaram piores ajustes, independente da função usada para modelar os efeitos aleatórios.

Após a escolha do melhor modelo utilizando polinômios segmentados do tipo B, foram inspecionados os componentes de variância estimados, a herdabilidade para todos os pontos ao longo da lactação e as correlações genéticas entre as produções de leite nos dias 6, 71, 137, 203 e 270 ajustadas pelo mesmo.

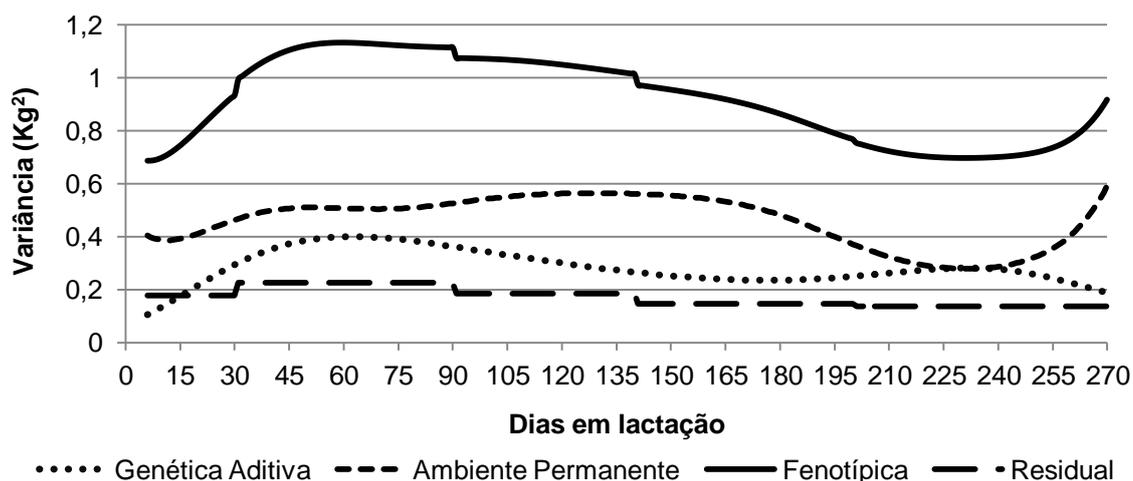


Figura 4 – Estimativas de variâncias obtidas pelo BSQ6H5.

De acordo com a Figura 4, foram encontrados maiores estimativas de variância genética aditiva e residual na fase do pico de produção, que ocorre por volta dos 40 dias após o parto. Uma das possíveis explicações seria o efeito de escala, já que nessa fase são observadas as maiores produções da lactação. A variância residual foi maior na fase do pico de produção e reduziu gradualmente até o período final.

As variâncias de ambiente permanente e fenotípica parecem estar subestimadas no início da lactação e superestimadas no final. Isso pode ter ocorrido devido ao menor número de dados disponíveis no início e principalmente no final da lactação, como consequência da proximidade do período seco e pelo fato de várias cabras já se encontrarem em período avançado da gestação.

O efeito aleatório de ambiente permanente variou ao longo da trajetória da lactação. Isso demonstra que ao ajustar coeficientes de (co) variâncias para este efeito, é possível separar mais eficientemente a variação de fatores genéticos de não genéticos, refletindo em estimativas de herdabilidade mais acuradas. As estimativas foram maiores no final da lactação. Melo et al. (2003), Cobuci et al. (2005) e Costa et al. (2008), em trabalhos com bovinos da raça Holandês, observaram maiores valores também no final e início da lactação, indicando que fatores ambientais exercem uma maior influência durante esses períodos. A Figura 5 mostra as estimativas de herdabilidade obtidas pelo modelo BSQ6H5.

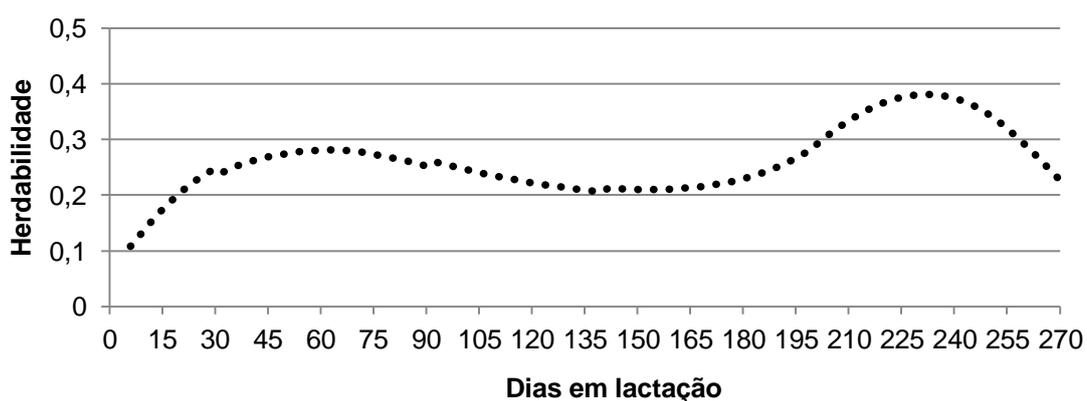


Figura 5 – Estimativas de herdabilidade para PLDC, obtidas pelo modelo BSQ6H5.

As estimativas de herdabilidade foram de menores magnitudes nos extremos da curva (Figura 5). Isso pode ser explicado pelas maiores

variâncias de ambiente permanente e pelas menores variâncias genética aditiva nesses estágios.

Na Tabela 7 são apresentadas as correlações genéticas para PLDC entre os dias 6, 71, 137, 203 e 270, obtidas pelo modelo BSQ6H5. As estimativas de correlações genéticas foram positivas e maiores entre controles leiteiros adjacentes e reduziram à medida que se aumentou o intervalo entre os controles leiteiros. Apesar das correlações genéticas entre produções mais distantes terem sido de menores magnitudes, todas as estimativas foram positivas indicando que a seleção dos animais para a produção de leite em um dia qualquer da lactação trará respostas positivas para todos os demais pontos da curva. Os resultados encontrados foram bem similares aos encontrados para o modelo LEG6H5.

Tabela 7 – Estimativas de correlações genéticas para PLDC's obtidas pelo modelo BSQ6H5.

PLDC	6	71	137	203	270
6		0,96	0,86	0,68	0,19
71			0,94	0,74	0,18
137				0,87	0,35
203					0,75
270					

c) Comparação entre modelos utilizando polinômios ortogonais de

Legendre e polinômios segmentados do tipo B

Após a escolha dos melhores modelos utilizando polinômios ortogonais de Legendre e polinômios segmentados do tipo B, foi feita uma comparação entre eles para a recomendação do melhor modelo a ser utilizado em avaliações genéticas de caprinos da raça Alpina. Na Tabela 8 são apresentados os dois modelos juntamente com os critérios utilizados para contrastá-los.

Tabela 8 – Melhores modelos, número de parâmetros (p), AIC, BIC e Log L.

Modelo	p	Log L	AIC	BIC
BSQ6H5	47	3223,592	-6353,184	-5988,488
LEG6H5	47	3179,123	-6264,246	-5899,548

Valores em negrito indicam o melhor modelo com base no AIC e BIC.

Em geral modelos mais parametrizados tendem a ser mais penalizados, uma vez que se buscam modelos mais parcimoniosos. Neste estudo, os dois melhores modelos possuem o mesmo número de parâmetros. Os critérios AIC e BIC penalizarão o número de parâmetros de forma similar, sendo assim apenas a comparação pelo Log L já seria adequada.

O modelo BSQ6H5 apresentou maior Log L do que o modelo LEG6H5, indicando melhor ajuste. Isto ocorreu possivelmente porque embora o número de parâmetros seja igual, é também um número alto, ou seja, ambos são modelos relativamente parametrizados, e nestes casos espera-se maiores problemas de multicolinearidade nos modelos utilizando polinômios ortogonais de Legendre do que aqueles que utilizam polinômios segmentados do tipo B, isto porque este reduz problemas de dependências das estimativas dos parâmetros do início com o final da regressão.

De acordo com as Figuras 2 e 4 (mostradas anteriormente), as variâncias fenotípica, genética aditiva, de ambiente permanente e residuais, obtidas pelos dois modelos (LEG6H5 e BSQ6H5) foram muito similares e apresentaram a mesma tendência ao longo da lactação. Inclusive o número de classes de variâncias residuais foi o mesmo para ambos os modelos (cinco).

Os dois modelos tenderam a superestimar a variância de ambiente permanente nos extremos da lactação, principalmente no final, possivelmente devido ao menor número de registros. Meyer (2005) afirmou que modelos de regressão aleatória ajustados por *B-splines* tendem a ser menos susceptíveis a problemas de estimativas erráticas nos extremos da curva, um problema frequentemente encontrado com polinômios ortogonais de Legendre.

Na Figura 6 são apresentadas as estimativas de herdabilidade obtidas pelos dois modelos. As estimativas obtidas pelo modelo LEG6H5 foram maiores ao longo de quase toda a lactação. Contudo, apresentaram a mesma tendência de variação ao longo da lactação, indicando que os períodos que possuem maior herdabilidade são coincidentes para os dois modelos. As estimativas de herdabilidade obtidas pelos dois modelos sugerem haver variabilidade genética suficiente para a obtenção de ganhos genéticos para a característica PLDC no rebanho estudado.

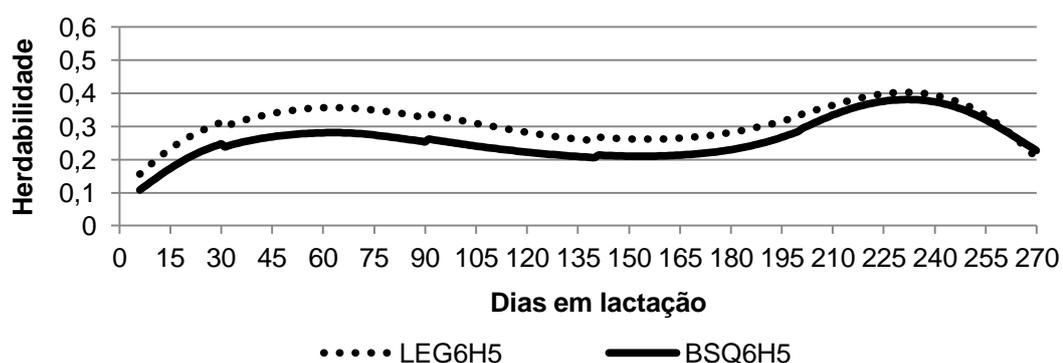


Figura 6 – Estimativas de herdabilidade para PLDC ao longo da lactação para os modelos LEG6H5 e BSQ6H5.

De acordo com os critérios AIC, BIC e Log L, o modelo que melhor se ajustou aos dados foi o modelo BSQ6H5. Além disso, este modelo contém segmentos individuais de polinômios de menores graus, o que permite um

melhor controle da influência global de observações individuais e tendem a proporcionar melhores estimativas. Sendo assim, conclui-se que para a avaliação genética de caprinos da raça Alpina, para a característica PLDC, deve se utilizar um modelo de regressão aleatória ajustado por polinômios segmentados do tipo B, de grau 2 (quadrático), para as curvas média da população, genética aditiva e de ambiente permanente, com cinco segmentos (seis nós) e cinco classes de variâncias residuais.

4.3. Raça Saanen

Na Tabela 9 encontram-se o número de observações, média e desvio padrão para a característica Produção de Leite no Dia do Controle (PLDC) de cabras da raça Saanen.

Os valores apresentados encontram-se dentro dos padrões aceitáveis para caprinos de aptidão leiteira. Apesar de a raça Saanen ser conhecida como a raça leiteira mais produtiva, no rebanho em estudo foram observadas produções menores que as observadas para a raça Alpina. Este resultado pode ser devido ao fato dos animais da raça Alpina adquiridos para o rebanho estudado serem de melhor qualidade genética e/ou à melhor adaptação da raça Alpina às condições ambientais da região. Soares Filho et al. (2001), no Distrito Federal encontraram média de PLDC de 2,34 kg/dia e Tholon et al. (2001) de 2,7 kg/dia, em rebanhos da região Sudeste. Isso mostra que patamares produtivos mais altos podem ser alcançados com o melhoramento genético do rebanho em estudo.

Tabela 9 – Número de registros (NR), média e desvio-padrão (DP) para cada produção de leite na semana do controle da primeira lactação de cabras da raça Saanen.

Semana	NR	Média (Kg)	DP	Semana	NR	Média (Kg)	DP
2	325	1,94	0,73	21	379	1,98	0,90
3	399	2,09	0,81	22	362	1,89	0,85
4	393	2,25	0,87	23	364	1,87	0,85
5	408	2,26	0,88	24	358	1,80	0,83
6	397	2,28	0,86	25	351	1,79	0,83
7	391	2,32	0,89	26	368	1,67	0,78
8	387	2,27	0,86	27	357	1,67	0,82
9	406	2,27	0,90	28	337	1,62	0,76
10	407	2,27	0,87	29	320	1,61	0,76
11	399	2,28	0,91	30	316	1,56	0,77
12	392	2,22	0,87	31	325	1,52	0,73
13	390	2,24	0,89	32	308	1,46	0,72
14	354	2,19	0,91	33	308	1,48	0,71
15	389	2,14	0,88	34	266	1,46	0,74
16	398	2,13	0,80	35	255	1,46	0,70
17	387	2,09	0,86	36	262	1,48	0,74
18	386	2,07	0,80	37	245	1,51	0,76
19	381	1,99	0,84	38	238	1,46	0,78
20	376	2,00	0,87	39	163	1,51	0,80

Na Figura 7 é apresentada a PLDC em função da semana de lactação. A produção de leite apresentou uma forma típica da curva de lactação de caprinos leiteiros, iniciando com 1,94 Kg no primeiro controle leiteiro, aumentando até 2,32 Kg na sétima semana de lactação após o parto e reduzindo gradualmente nos controles leiteiros subsequentes até 1,51 Kg na 39ª semana.

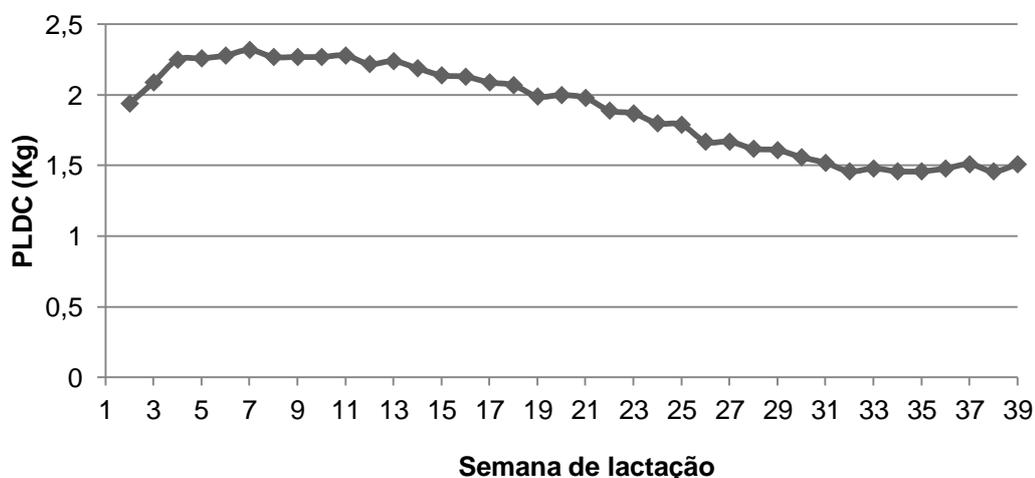


Figura 7 – Produção de leite ao longo da lactação de cabras da raça Saanen.

4.3.1. Ajuste e comparação de modelos

a) Polinômios Ortogonais de Legendre

O emprego de modelos de regressão aleatória por meio de funções polinomiais de Legendre requer, inicialmente, a definição da ordem mais apropriada para a regressão média da população e para os efeitos aleatórios considerados no modelo de análise (genético aditivo e de ambiente permanente) e a avaliação da necessidade de considerar a heterogeneidade de variância residual.

Primeiramente serão apresentados na Tabela 10 os resultados dos modelos de regressão aleatória ajustados por polinômios ortogonais de Legendre considerando homogeneidade de variâncias residuais, além dos critérios utilizados para contrastá-los.

Os valores de Log L, AIC e BIC melhoraram com o aumento no número de parâmetros e de acordo com o TRV foram encontradas diferenças significativas ($P < 0,01$) entre todos os modelos reduzidos e o modelo completo (Tabela 10).

Tabela 10 – Modelos, número de parâmetros (p), AIC, BIC, Log L e TRV para os polinômios ortogonais de Legendre, considerando homogeneidade de variâncias residuais.

Modelo*	p	Log L	AIC	BIC	TRV
LEG3H1	13	3423,000	-6820,000	-6722,798	**
LEG4H1	21	3645,136	-7248,272	-7091,254	**
LEG5H1	31	3787,552	-7513,104	-7281,318	**
LEG6H1	43	3916,180	-7746,360	-7424,856	completo

** = significativo a 1%; *Modelo: LEGxHy, em que x corresponde à ordem dos polinômios para a curva fixa e para os efeitos aleatórios genético aditivo e de ambiente permanente e y, ao número de classes de variâncias residuais; valores em negrito indicam o melhor modelo com base no AIC e BIC.

Assim como nas análises da raça Alpina, os modelos mais parametrizados apresentaram maior dificuldade de convergência. Isto possivelmente ocorreu devido ao aumento da multicolinearidade nas

matrizes de componentes de variância genética aditiva e de ambiente permanente, dificultando a convergência de análises REML e prejudicando a qualidade das estimativas dos componentes.

Dentre os modelos que consideraram homogeneidade de variâncias residuais, o modelo LEG6H1 foi o que melhor se ajustou aos dados segundo os critérios utilizados para compará-los. Entretanto, também foi avaliada a necessidade de considerar a heterogeneidade de variância residual.

Tabela 11 – Modelos, número de parâmetros (p), Log L, AIC, BIC e TRV para os polinômios ortogonais de Legendre, considerando heterogeneidade de variâncias residuais.

Modelo*	p	Log L	AIC	BIC	TRV
LEG6H1	43	3916,180	-7746,360	-7424,856	**
LEG6H2	44	4011,849	-7935,698	-7606,718	**
LEG6H3	45	3999,807	-7909,614	-7573,156	**
LEG6H4	46	4005,039	-7918,078	-7574,144	**
LEG6H5	47	4012,231	-7930,462	-7579,050	**
LEG6H6	48	4024,553	-7953,106	-7594,218	completo

** = significativo a 1%; *Modelo: LEGxHy, em que x corresponde à ordem dos polinômios para a curva fixa e para os efeitos aleatórios genético aditivo e de ambiente permanente e y, ao número de classes de variâncias residuais; valores em negrito indicam o melhor modelo com base no AIC e BIC.

De acordo com os resultados apresentados na Tabela 11, verificou-se que o modelo que desconsidera a heterogeneidade de variância residual obteve o pior ajuste dentre todos os modelos avaliados. De acordo com Menezes (2008), essa heterogeneidade pode estar associada a diversos fatores, como período de gestação, condições do parto e duração do período seco. À medida que se aumentou o número de classes do resíduo, houve melhoria do ajuste, como indicado por todos os critérios utilizados. Menezes (2008), em caprinos da raça Saanen, também verificou essa tendência de melhoria do ajuste com o aumento no número de classes de variâncias residuais.

Geralmente os modelos mais parcimoniosos são preferidos, já que um aumento no número de parâmetros a serem estimados pode ocasionar dificuldades de convergência, especialmente em banco de dados com grande volume de observações. A adição de uma nova classe ocasiona um aumento no número de parâmetros bem inferior do que quando se aumenta a ordem das regressões genética aditiva e de ambiente permanente. Conforme observado na Tabela 10, quando se aumentou apenas a ordem dos polinômios, o número de parâmetros variou de 13 a 43, sendo o número de parâmetros acrescidos dependentes da ordem subsequente. De acordo com a Tabela 11, quando se variou apenas o número de classes, houve um aumento de apenas um parâmetro por classe. Apesar disso, o modelo que melhor se ajustou aos dados segundo os critérios utilizados foi o modelo LEG6H6.

Após a escolha do melhor modelo utilizando polinômios ortogonais de Legendre, foram inspecionados os componentes de variância estimados pelo modelo, a herdabilidade para todos os pontos da lactação e as correlações genéticas entre as produções de leite nos dias 6, 71, 137, 203 e 270. A Figura 8 apresenta as variâncias, genética aditiva, de ambiente permanente, residual e fenotípica, ajustadas pelo modelo LEG6H6.

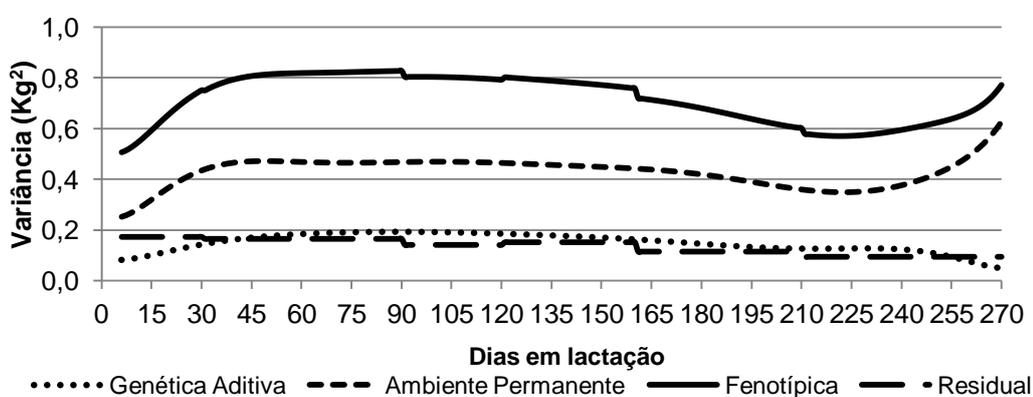


Figura 8 – Estimativas de variâncias obtidas pelo modelo LEG6H6.

A Figura 8 mostra que as variâncias, genética aditiva e de ambiente permanente, foram bem constantes ao longo da lactação, com exceção da fase inicial e final. No final da lactação a variância genética aditiva diminuiu. Os resultados encontrados para a raça Alpina, tanto para o modelo LEG6H5 quanto para o modelo BSQ6H5 também indicaram menores estimativas de variâncias aditivas no final da lactação. Entretanto, Silva (2011) encontrou maiores estimativas no final da lactação e Bignardi (2010), em trabalho com bovinos, observou maiores estimativas nas duas primeiras semanas de lactação.

A variância residual foi maior no início da lactação e reduziu gradualmente até o período final. Isso mostra que houve maiores influências do ambiente nesse período e ressalta a importância da consideração da heterogeneidade de variância residual, o que contribuiu para uma melhor partição da variação fenotípica.

De acordo com a Figura 8, parece haver problemas nas estimativas de variâncias nos extremos da curva. Isso pode ter ocorrido devido ao menor número de dados disponíveis nesses estágios da lactação, principalmente no final da lactação em que houve uma redução acentuada no número de registros. As estimativas das variâncias nas análises para a raça Alpina também seguiram a mesma tendência, contudo, foram de menores magnitudes.

A Figura 9 mostra que as estimativas de herdabilidade variaram de 0,06 a 0,24 ao longo da lactação. Resultados similares foram obtidos por Menezes (2008) em caprinos da raça Saanen, que encontrou estimativas de herdabilidade variando de 0,07 a 0,25. Assis et al. (2006) e Breda et al.

(2006), obtiveram estimativas de herdabilidade máximas e mínimas iguais a 0,32 e 0,09; 0,51 e 0,12, respectivamente, para cabras da raça Alpina.

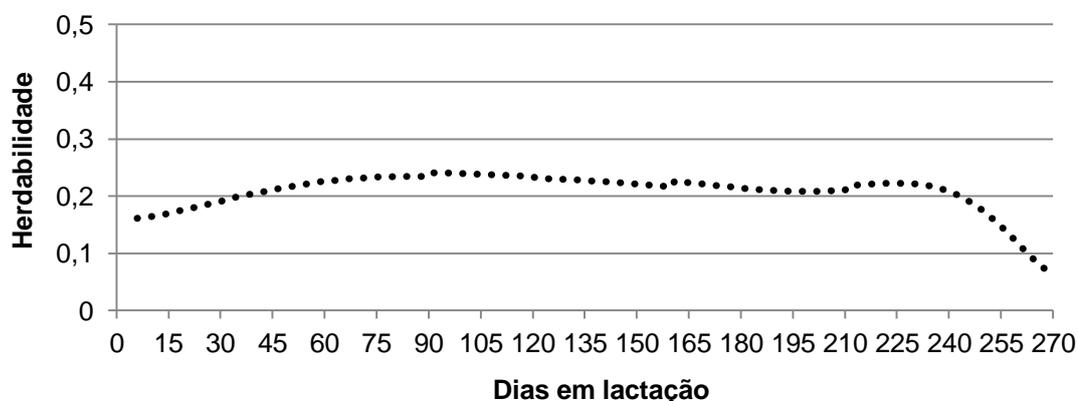


Figura 9 – Estimativas de herdabilidade para PLDC obtidas pelo modelo LEG6H6.

As estimativas de herdabilidade foram menores nos extremos da lactação. Isto pode ter ocorrido devido à maior influência dos efeitos ambientais nessas fases. No início da lactação, a cabra sofre um estresse pós-parto devido ao balanço energético negativo e no final da lactação, muitas delas já estão em estágio avançado da gestação. Além disso, conforme já mencionado, nesses dois extremos há um menor número de registros, o que também pode ocasionar erros nas estimativas das variâncias e conseqüentemente nas estimativas de herdabilidade e correlações genéticas, principalmente por que a perda de dados nestes períodos leva a uma extrapolação nas estimativas de ambiente permanente para alguns indivíduos. Resultados semelhantes foram obtidos por Menezes (2008).

Foi observada uma queda abrupta de herdabilidade após os 250 dias de lactação, não condizente com a realidade biológica, onde se esperam alterações mais suaves nos parâmetros. Esta discrepância pode ser proveniente da incapacidade do modelo de prover ajustes adequados em

regiões com menor número de dados e conforme pode ser observado na Tabela 9, houve um decréscimo considerável no número de registros nas duas últimas semanas principalmente.

Tabela 12 – Estimativas de correlações genéticas para PLDC's obtidas pelo modelo LEG6H6.

PLDC	6	71	137	203	270
6		0,85	0,77	0,70	0,61
71			0,89	0,85	0,83
137				0,94	0,74
203					0,82
270					

As correlações genéticas aditivas entre os dias analisados variaram de 0,61 a 0,94. Os valores altos e positivos indicam que a seleção para aumentar a produção de leite no início da lactação, tende a proporcionar um aumento nas produções de leite subsequentes. Esses valores são maiores que os obtidos para a raça Alpina e estão mais próximos dos encontrados por Silva (2011), que encontrou valores próximos da unidade. Bignardi et al. (2009), em trabalho com bovinos, também encontraram correlações genéticas entre PLDC's mais altas quanto mais próximos os controles leiteiros.

b) **Polinômios segmentados do tipo B (B-splines)**

Após o estudo dos polinômios ortogonais de Legendre, avaliou-se o ajuste de modelos de regressão aleatória utilizando polinômios segmentados do tipo B. A Tabela 13 apresenta os modelos que foram testados considerando homogeneidade de variâncias residuais.

Tabela 13 – Modelos, número de parâmetros (p), AIC, BIC, Log L e TRV para os polinômios segmentados do tipo B, considerando homogeneidade de variâncias residuais.

Modelo*	P	Log L	AIC	BIC	TRV
Linear					
BSL3H1	13	3330,606	-6635,212	-6538,010	**
BSL4H1	21	3644,481	-7246,962	-7089,944	**
BSL5H1	31	3759,933	-7457,866	-7226,080	**
BSL6H1	43	3781,953	-7477,906	-7152,402	completo
Quadrático					
BSQ3H1	13	3400,275	-6774,550	-6677,348	**
BSQ4H1	21	3666,052	-7290,104	-7133,088	**
BSQ5H1	31	3791,584	-7521,168	-7289,382	**
BSQ6H1	43	3907,422	-7728,844	-7407,340	completo
Cúbico					
BSC4H1	21	3652,308	-7262,616	-7105,598	**
BSC5H1	31	3787,087	-7512,174	-7280,390	**
BSC6H1	43	3924,565	-7763,130	-7441,626	completo

** = significativo a 1%; *Modelo: BSXkHy, em que X = L, Q e C são os graus do segmento polinomial linear, quadrático e cúbico, respectivamente, k corresponde ao número de nós para a curva média da população e para os efeitos genético aditivo e de ambiente permanente e y, ao número de classes de variâncias residuais; valores em negrito indicam o melhor modelo com base no AIC e BIC.

O software WOMBAT não permite a avaliação de modelos BSC3 (três nós), pois para *B-splines* do tipo cúbico, espera-se que ao menos um parâmetro influencie a variável dependente em pelo menos três segmentos, para tal seriam necessários ao menos quatro nós.

De acordo com os resultados apresentados na Tabela 13, independentemente da ordem da função *B-spline* (linear, quadrática ou cúbica) empregada, quando se aumentou o número de nós, tanto o AIC como o BIC apresentaram valores inferiores, indicando melhores ajustes. Considerando homogeneidade de variâncias residuais, o modelo que melhor se ajustou aos dados foi o modelo BSC6H1.

Devido ao fato de o TRV não permitir comparação entre os melhores modelos de cada grau (BSL6H1, BSQ6H1 e BSC6H1), foram avaliados

números diferentes de classes de variâncias residuais para os polinômios segmentados de grau 1, 2 e 3. Na Tabela 14 estão apresentados os modelos para os polinômios segmentados considerando heterogeneidade de variância residual.

Tabela 14 – Modelos, número de parâmetros (p), AIC, BIC, Log L e TRV para os polinômios segmentados do tipo B, considerando heterogeneidade de variâncias residuais.

Modelo*	p	Log L	AIC	BIC	TRV
Linear					
BSL6H1	43	3781,953	-7477,906	-7152,402	**
BSL6H2	44	3845,138	-7602,276	-7273,294	**
BSL6H3	45	3886,804	-7683,608	-7347,148	**
BSL6H4	46	3934,387	-7776,774	-7432,840	**
BSL6H5	47	3940,128	-7786,256	-7434,814	**
BSL6H6	48	3964,152	-7832,304	-7473,414	completo
Quadrático					
BSQ6H1	43	3907,422	-7728,844	-7407,340	**
BSQ6H2	44	3995,399	-7902,798	-7573,818	**
BSQ6H3	45	3988,668	-7887,336	-7550,878	**
BSQ6H4	46	3996,268	-7900,536	-7556,660	**
BSQ6H5	47	4004,705	-7915,410	-7563,998	**
BSQ6H6	48	4015,189	-7934,378	-7575,488	completo
Cúbico					
BSC6H1	43	3924,565	-7763,130	-7441,626	**
BSC6H2	44	3983,939	-7879,860	-7550,880	**
BSC6H3	45	4006,189	-7922,378	-7585,920	**
BSC6H4	46	4008,098	-7924,196	-7580,260	**
BSC6H5	47	4021,751	-7949,502	-7598,090	completo

** = significativo a 1%; ns = não-significativo; valores em negrito indicam o melhor modelo com base no AIC e BIC; *Modelo: BSXkHy, em que X = L, Q e C são os graus do segmento polinomial linear, quadrático e cúbico, respectivamente, k corresponde ao número de nós para a curva média da população e para os efeitos genético aditivo e de ambiente permanente e y, ao número de classes de variâncias residuais.

O modelo BSC6H6 não foi considerado, pois apresentou problemas na estimação dos componentes de variância e de convergência pelo algoritmo AI-REML, mesmo quando as análises foram reiniciadas com diversos valores para os parâmetros.

De acordo com os critérios utilizados, o modelo que melhor se ajustou aos dados foi o modelo BSC6H5. DeGroot et al. (2007) utilizando registros de três lactações de bovinos da raça Holandês, concluíram que uma função *spline* cúbica foi eficiente para descrever a produção de leite no decorrer da lactação.

Após a escolha do melhor modelo utilizando polinômios segmentados do tipo B, foram inspecionados os componentes de variância, a herdabilidade para todos os pontos da curva de lactação e as correlações genéticas entre as produções de leite nos dias 6, 71, 137, 203 e 270 ajustadas pelo mesmo.

De acordo com a Figura 10, as estimativas de variância foram similares às encontradas pelo modelo LEG6H6. As variâncias de ambiente permanente e genética aditiva foram mais constantes ao longo da lactação, com exceção dos extremos da curva, onde as variâncias de ambiente permanente parecem estar subestimadas no início da lactação e superestimadas no final. Isso pode ter ocorrido devido ao menor número de registros nessas fases.

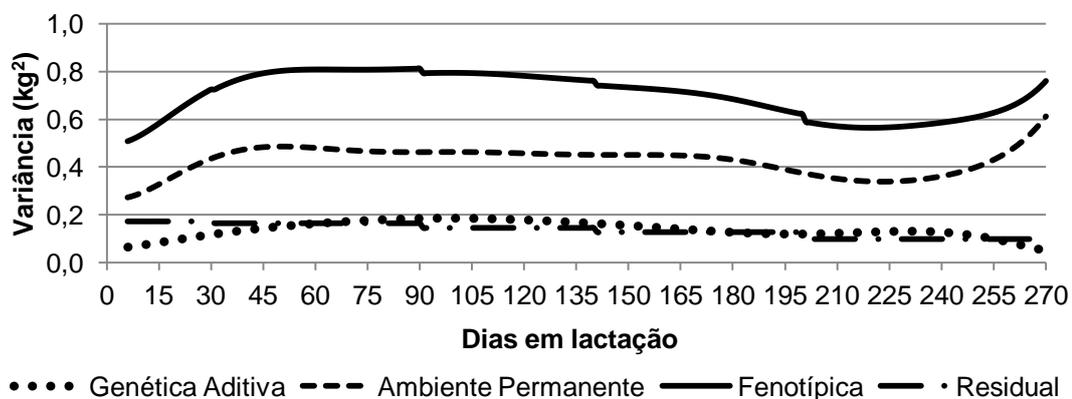


Figura 10 – Estimativas de variâncias obtidas pelo modelo BSC6H5.

A variância residual seguiu a mesma tendência das estimativas obtidas pelo modelo LEG6H6, sendo encontrados maiores valores no início da lactação, reduzindo gradualmente até o período final. Isso mostra que houve maiores influências do ambiente nesse período e ressalta a importância da consideração da heterogeneidade de variância residual, o que contribuiu para uma melhor partição da variação fenotípica.

As mais altas variâncias de ambiente permanente no final da lactação estimadas com o modelo BSC6H5 também foram observadas por Melo et al. (2003), Cobuci et al. (2005) e Costa et al. (2008) para bovinos leiteiros.

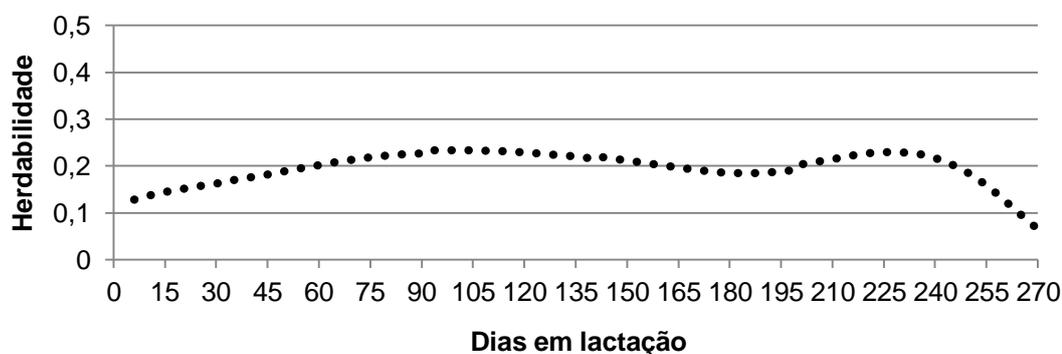


Figura 11 – Estimativas de herdabilidade obtidas pelo modelo BSC6H5.

De acordo com a Figura 11, as estimativas de herdabilidade foram de menores magnitudes nos extremos da curva. Isso pode ser explicado pelas maiores variâncias de ambiente permanente e por menor variância genética aditiva nesses estágios. Silvestre et al. (2006), em trabalho com bovinos, utilizando um modelo de regressão aleatória empregando uma função *B-spline* cúbica e Costa et al. (2008), em estudos realizados usando MRA utilizando polinômios ortogonais de Legendre na raça Holandês no Brasil, reportaram que os valores de herdabilidade também foram maiores na fase intermediária da lactação.

Na Tabela 15 são apresentadas as correlações genéticas para PLDC entre os dias 6, 71, 137, 203 e 270, obtidas pelo modelo BSC6H5. As correlações genéticas aditivas entre os dias analisados variaram de 0,64 a 0,94. Os resultados foram similares aos obtidos pelo modelo LEG6H6. Os valores altos e positivos indicam que a seleção para aumentar a produção de leite no início da lactação, tende a proporcionar um aumento nas produções de leite subsequentes. Esses valores são maiores que os obtidos para a raça Alpina (neste estudo) e estão mais próximos dos encontrados por Silva (2011), que encontrou valores próximos da unidade.

Tabela 15 – Correlações genéticas entre algumas PLDC's obtidas pelo modelo BSC6H5.

PLDC	6	71	137	203	270
6		0,86	0,74	0,69	0,64
71			0,84	0,83	0,66
137				0,94	0,72
203					0,79
270					

c) Comparação entre os modelos utilizando polinômios ortogonais de Legendre e polinômios segmentados do tipo B

Após a escolha dos melhores modelos utilizando polinômios ortogonais de Legendre e polinômios segmentados do tipo B, foi feita uma comparação entre eles para a recomendação do melhor modelo a ser utilizado em avaliações genéticas de caprinos da raça Saanen. A Tabela 16 apresenta os dois modelos juntamente com os critérios utilizados para contrastá-los.

Tabela 16 – Melhores modelos, número de parâmetros (p), Log L, AIC e BIC.

Modelo	p	Log L	AIC	BIC
BSC6H5	47	4021,751	-7949,502	-7598,090
LEG6H6	48	4024,553	-7953,106	-7594,218

Geralmente modelos mais parametrizados tendem a obter maiores valores de Log L, já que esse critério não penaliza o número de parâmetros. Como o modelo BSC6H5 possui um parâmetro a menos, obteve um melhor valor de BIC, já que este critério penaliza de forma mais rigorosa modelos mais parametrizados.

De acordo com as Figuras 8 e 10 (mostradas anteriormente), as variâncias fenotípica, genética aditiva, de ambiente permanente e residuais, obtidas pelos dois modelos (LEG6H6 e BSC6H5) foram muito similares e mostraram a mesma tendência ao longo da lactação. Além disso, as correlações genéticas e as estimativas de herdabilidade foram muito semelhantes, como mostrado na Figura 12. Sendo que os dois modelos tenderam a superestimar a variância de ambiente permanente nos extremos da lactação, principalmente no final, possivelmente devido ao menor número de registros.

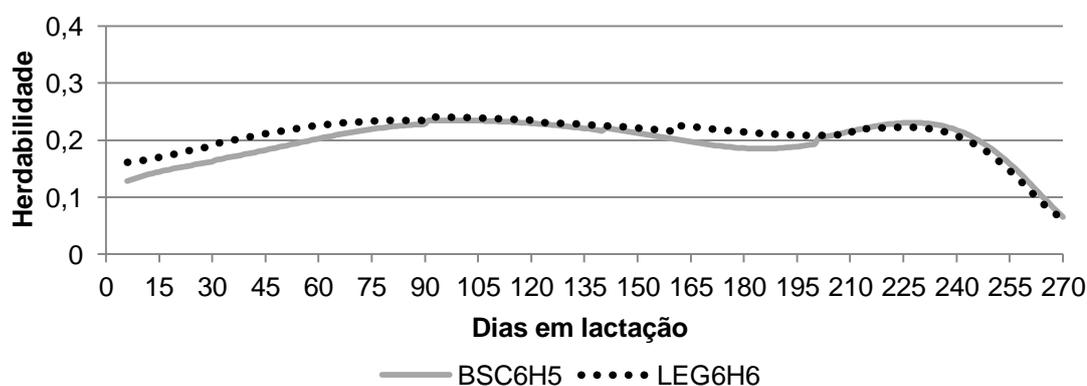


Figura 12 – Estimativas de herdabilidade para os dois melhores modelos.

As variâncias implausíveis no final da lactação foram observadas para os dois modelos. Este resultado sugere que esse efeito não é causado apenas pelo tipo de função utilizada, mas pode ser causado também pelo menor número de informações nos extremos da lactação e fatores não

considerados como tratamentos preferenciais e variações de rebanho na forma da curva de lactação e efeitos de gestação.

As diferenças entre os modelos LEG6H6 e BSC6H5 foram pequenas, indicando que qualquer um dos dois modelos pode ser utilizado para avaliação genética da produção de leite de cabras da raça Saanen. Entretanto, sugere-se a utilização do modelo BSC6H5 por serem polinômios segmentados de menor grau, o que permite um melhor controle da influência global de observações e tendem a proporcionar melhores estimativas nos extremos da curva; é menos parametrizado, o que é vantajoso principalmente para avaliações genéticas com maior número de registros; e possui um menor número de classes de variâncias residuais.

Sendo assim, concluímos que para a avaliação genética da característica PLDC em caprinos da raça Saanen, sugere-se utilizar um modelo de regressão aleatória ajustado por polinômios segmentados do tipo B, de grau 3 (cúbico), com cinco segmentos (seis nós) e cinco classes de variâncias residuais.

Neste trabalho, observa-se que foram indicados modelos diferentes para as raças Alpina e Saanen, apesar de todos os animais, das diferentes raças estudadas, serem criados em conjunto, sob o mesmo manejo. Esses resultados corroboram com Oliveira (2008), que avaliou distintas funções polinomiais com diferentes ordens para o melhor ajuste das regressões fixas e aleatórias da trajetória de crescimento de ovinos das raças Santa Inês, Poll Dorset e Somalis Brasileira. A autora ressaltou que apesar de todos os animais, das diferentes raças estudadas, serem criados em conjunto, sob o mesmo manejo, diferentes modelos foram necessários para o ajuste da

trajetória de crescimento das diferentes raças, principalmente no que se refere à parte fixa desta trajetória.

5 – CONCLUSÃO

O uso de polinômios segmentados do tipo B com heterogeneidade de variância residual é recomendado para a avaliação genética da produção de leite de caprinos, cabendo algumas particularidades para diferentes raças.

6 – 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.
- ALBUQUERQUE, L.G. Regressão aleatória: nova tecnologia pode melhorar a qualidade das avaliações genéticas. In: V Simpósio Brasileiro de Melhoramento Animal, 2004, Pirassununga. **Anais...** Pirassununga: SBMA, 2004.
- ALBUQUERQUE, L.G.; PEREIRA, C.S. Evolução dos modelos de avaliação genética e novos desafios. In: 43ª Reunião Anual da SBZ. **Anais...** João Pessoa, PB, 2006.
- ARAÚJO, C.V.; BITTENCOURT, T.C.B.S. C.; ARAÚJO, S.I. et al. Heterogeneidade de variâncias na avaliação genética para peso ao ano em bovinos de corte da raça Nelore. In: VIII Simpósio Brasileiro de Melhoramento Animal - Melhoramento Animal: UMA VISÃO CRÍTICA. **Anais...**, Maringá, PR. 01 e 02 de julho de 2010.
- ARAÚJO, C.V.; TORRES, R.A.; COSTA, C.N.; et al. Uso de funções ortogonais para descrever a produção de leite no dia de controle por meio de modelos de regressão aleatória. **Revista Brasileira de Zootecnia**, v.35, n.3, p.967-974, 2006.
- ASSIS, G.M.L.; ALBUQUERQUE, L.G. ; SARMENTO, J.L.R. et al. Estimação de parâmetros genéticos em caprinos leiteiros por meio da análise de regressão aleatória, utilizando a amostragem de Gibbs. **Revista Brasileira de Zootecnia**, v.35, n.3, p.706-714, 2006.
- BIGNARDI, A.B. EL FARO, L. ALBUQUERQUE, L.G. et al. Modelos de regressão aleatória empregando funções B-splines para modelar a produção de leite de vacas da raça Holandesa. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 45., 2008, Lavras. **Anais...** Lavras: SBZ, 2008. (CD-ROM).
- BIGNARDI, A.B. **Modelos para estimação de componentes de (co) variância para produção de leite no dia do controle de vacas da Raça Holandesa**. Jaboticabal, UNESP, 2010. 57p. Tese (Doutorado) – Universidade Estadual Paulista, 2010.

- BIGNARDI, A.B., EL FARO, L., TORRES JÚNIOR, R.A.A., et al. Random regression models using different functions to model test-day milk yield of Brazilian Holstein cows. **Genetic and Molecular Research**, v.10, n.4, p. 3565-3575, 2011.
- BISPO, C.H.S. **Fertilidade do sêmen caprino resfriado ou congelado em diferentes concentrações de gema-de-ovo no diluente**. Viçosa; UFV, 2009. 110p. Tese (Doutorado) – Universidade Federal de Viçosa, 2009.
- BOHMANOVA, J., MIGLIOR, F., JAMROZIK, J. et al. Comparison of random regression models with Legendre polynomials and linear splines for production traits and somatic cell score of Canadian Holstein cows. **Journal of Dairy Science**, v.91, p.3627–3638, 2008.
- BREDA, F.C.; ALBUQUERQUE, L.G.; YAMAKI, M. et al. Estimação de parâmetros genéticos para produção de leite de cabras da raça Alpina. **Revista Brasileira de Zootecnia**, v.35, n.2, p.396-404, 2006.
- CHAVES, L.C.S. **Avaliação da persistência da lactação em vacas Girolando utilizando modelos de regressão aleatória**. Viçosa; UFV, 2005. 86p. Dissertação (Mestrado) – Universidade Federal de Viçosa, 2005.
- CHAVES, L.C.S. **Avaliação genética em bubalinos leiteiros utilizando modelos de regressão aleatória**. Viçosa; UFV, 2009. 77p. Tese (Doutorado) – Universidade Federal de Viçosa, 2009.
- COBUCI, J.A. **Uso de modelos de regressão aleatória na avaliação da persistência na lactação de animais da raça Holandesa**. Viçosa, UFV, 2002. 99p. Tese (Doutorado) – Universidade Federal de Viçosa, 2002.
- COBUCI, J. A.; COSTA, C.N.; TEIXEIRA, N.M.; et al. Estimativa de herdabilidade para produção de leite no dia do controle de vacas da raça Holandesa utilizando regressão aleatória. In: Anais do V Simpósio da Sociedade Brasileira de Melhoramento Animal. **Anais...**, Pirassununga-SP, 2004.
- COBUCI, J.A.; EUCLYDES, R.F.; LOPES, P.S. et al. Estimation of genetic parameters for test-day milk in Holstein cows using a random

- regression model. **Genetics and Molecular Biology**, v.28, n°1, p.75-83, 2005.
- COSTA, C.N.; MELO, C.N.R.; PACHER, I.U. et al. Genetic parameters for test day milk yield of first lactation Holstein cows estimated by random regression using Legendre polynomials. **Revista Brasileira de Zootecnia**, v.37, p.602-608, 2008.
- DE BOOR, C. **A practical guide to splines**. Springer Series in Applied Mathematics, Springer Verlag, New York, 2nd ed., 2001.
- DeGROOT, B.J.; KEOWN, J.F.; VAN VLECK, L.D. et al. Estimates of genetic parameters for Holstein cows for test-day yield traits with a random regression cubic spline model. **Genetic and Molecular Research**, v.6, n.2, p. 434-444, 2007.
- EL FARO, L.; ALBUQUERQUE, L.G. Utilização de Modelos de Regressão Aleatória para Produção de Leite no Dia do Controle, com Diferentes Estruturas de Variância Residuais. **Revista Brasileira de Zootecnia**, v.32, p.1104-1113, 2003.
- FREITAS, L.S.; SILVA, M.A.; VERNEQUE, R.S.; et al. Comparação de modelos de regressão aleatória, para avaliação genética da produção de leite no dia do controle de vacas da Guzerá. In: VII Simpósio Brasileiro de Melhoramento Animal. **Anais...**, São Carlos-SP, 2008.
- GONÇALVES, H.C.; SILVA, M.A.; WECHSLER, F.S. et al. Parâmetros e tendência genética da produção de leite de cabra no Brasil. **Revista Brasileira de Zootecnia**, v.31, n.6, p.2204- 2208, 2002.
- GUIMARÃES, V. P.; RODRIGUES, M.T.; SARMENTO, J.L.R. et al. Utilização de funções matemáticas no estudo da curva de lactação em caprinos. **Revista Brasileira de Zootecnia**, v.35, n.2, p.535-543, 2006.
- HENDERSON JR., C.R., 1982. Analysis of covariance in the mixed model: higher level, non homogeneous, and random regressions. **Biometrics**, 38, 623-640.
- HUISMAN, A.E., VEERKAMP, R.F., VAN ARENDONK, J.A. Genetic parameters for various random regression models to describe the

- weight data of pigs. **Journal of Animal Science**, v.80, p.575-582, 2002.
- IBGE – Instituto Brasileiro de Geografia e Estatística: Informações Estatísticas – Economia – Pecuária. Disponível em: http://www.ibge.gov.br/home/estatistica/economia/ppm/2010/tabelas_pdf/tab01.pdf. Acessado em: 08 de maio de 2012.
- JACOPINI, L.A.; MARTINS, E.N.; LINO, D.A. et al. Comportamento da produção de leite de cabras Saanen e mestiças Boer - Saanen. In: Anais do VIII Simpósio Brasileiro de Melhoramento Animal. **Anais...**, Maringá-PR, 2010.
- JAMROZIK, J.; SCHAEFFER, L.R. Estimates of genetic parameters for a test day model with random regressions for production of first lactation Holsteins. **Journal Dairy Science**, v.80, p.762-770, 1997.
- JORGE JÚNIOR, J., DIAS, L.T., ALBUQUERQUE, L.G. Fatores de correção de escores visuais de conformação, precocidade e musculatura, à desmama, para idade da vaca ao parto, data juliana de nascimento e idade à desmama em bovinos da raça Nelore. **Revista Brasileira de Zootecnia**, v.33, n.6, p. 2044-2053, 2004.
- KEELE, L.J. **Semiparametric regression for the social sciences**. John Wiley and Sons, 1ª edição, 2008.
- LÓPEZ-ROMERO, P., R. REKAYA, AND M. J. CARABANO. Bayesian comparison of test-day models under different assumptions of heterogeneity for the residual variance: The change point technique versus arbitrary intervals. **J. Anim. Breed. Genet.** 121:14–25, 2004.
- MALHADO, C.H.M., RAMOS, A.A., CARNEIRO, P.L.S. et al. Parâmetros e tendências da produção de leite em bubalinos da raça Murrah no Brasil. **Revista Brasileira de Zootecnia**, v.36, n.2, p. 376-379, 2007.
- MENEZES, G. R.O.; BRITO, L.F.; MELO, A.L.P. et al. Comparação de modelos de regressão aleatória utilizados na avaliação da produção de leite em cabras da raça Saanen. Anais do VII Simpósio Brasileiro de Melhoramento Animal. **Anais...**, São Carlos-SP, 2008.
- MENEZES, G.R.O, **Avaliação da produção de leite em cabras da raça Saanen utilizando modelos de regressão aleatória**. Viçosa, UFV,

2008. 76p. Dissertação (Mestrado) – Universidade Federal de Viçosa, 2008.
- MENEZES, G.R.O.; CAETANO, G.C.; BRITO, L.F. et al. 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. **Anais...**, Maringá, PR. 01 e 02 de julho de 2010(a).
- MENEZES, G.R.O. **Uso de polinômios segmentados na modelagem de dados longitudinais de ponderal em bovinos da raça Tabapuã.** Viçosa, UFV, 2010. 99p. Tese (Doutorado) – Universidade Federal de Viçosa, 2010(b).
- 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.
- MEYER, K. Random regression analyses using B-splines to model growth of Australian Angus cattle. **Genetic Selection Evolution**, v.37, p. 473–500, 2005.
- 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.
- MUÑOZ-BERROCAL, M., TONHATI, H., CERÓN-MUÑOZ, M.F. et al. Uso de polinômios segmentados para o estudo da curva de lactação em búfalos Murrah e seus mestiços em sistema de criação extensivo, no estado de São Paulo. **Archivos Latinoamericanos de Produccion Animal**, v.13, n.1, p.24-29, 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, K.A.P. **Avaliação Genética de parte da trajetória de crescimento em ovinos das raças Santa Inês, Poll Dorset e Somalis Brasileira utilizando Modelos de Regressão Aleatória.** Fortaleza, UFCE, 2008. 69p. Dissertação (Mestrado) – Universidade Federal do Ceará, 2008.

- OLIVEIRA, D.C., **Funções *splines* para estudo de curvas de crescimento em ovinos cruzados**. Viçosa, UFV, 2011. 68p. Dissertação (Mestrado) – Universidade Federal de Viçosa, 2011.
- PAULSON, D.S. **Handbook of regression and modeling – Application for the clinical and pharmaceutical industries**. Chapman e Hall / CRC Biostatistics Series, 2007.
- PELICIONI, L.C., PASCOA, L., MUNIZ, C.M.S., et al. Efeito da idade da vaca ao parto e da data juliana de nascimento sobre características pré-desmama de bezerros da raça Gir. **Revista Brasileira de Zootecnia**, v.31, n.1, p. 61-70, 2002.
- PEREIRA, R.J. **Medidas de persistência da produção de leite em rebanhos Gir Leiteiro usando modelos de regressão aleatória**. Viçosa, UFV, 2009. 68p. Dissertação (Mestrado) – Universidade Federal de Viçosa, 2009.
- PEREIRA, R.J.; EL FARO, L.; VERNEQUE, R.S. et al. Modelos de regressão aleatória empregando B-spline para a produção de leite no dia do controle em bovinos Gir Leiteiro. In: Anais do VIII Simpósio da Sociedade Brasileira de Melhoramento Animal. **Anais...**, Maringá-PR, 2010.
- REGAZZI, A.J. Teste para avaliar a igualdade de parâmetros e a identidade de modelos de regressão não linear. **Revista Ceres**, v.50, n.287, p.9-26, 2003.
- REIS FILHO, J.C., **Alternativas para avaliação genética na raça Gir Leiteiro**. Viçosa, UFV, 2009. 85p. Tese (Doutorado) – Universidade Federal de Viçosa, 2009.
- REKAYA, R.; CARABAÑO, M.J.; TORO, M.A. Random regression in animal models for test-day production in dairy cattle. Proceedings of the fifth world congress on genetics applied to livestock production. **Proceedings...** Guelph, v.18, p.443-446. 1999.
- RIBEIRO, S. D. A. Caprinocultura: criação racional de caprinos. São Paulo: [s.n.], 1997. 311 p
- RICE, J.A., WU C.O., Non parametric mixed effects models for unequally sampled noisy curves. **Biometrics**, v.57, p.253-259, 2001.

- SÁNCHEZ, J.P., MISZTAL, I., AGUILAR, I. et al. Genetic evaluation of growth in a multibreed beef cattle population using random regression-linear spline models. **Journal of Animal Science**, v.86, p.267-277, 2008.
- SCHAEFFER, L.R. DEKKERS, J.C.M. Random regression in animal models for test day production in dairy cattle. In: Proceedings 5th world congress genetic applied livestock production. **Proceedings...** Guelph, ON, Canada, p.443-446, 1994.
- SCHAEFFER, L.R.; JAMROZIK, J.; KISTEMAKER, G.J. et al. Experience with a test day model. **Journal of Dairy Science**, v.83, p.1135-1144, 2000.
- SCHAEFFER, L.R. 2003. Quantitative genetics and animal models (course notes).
- SCHENKEL, F.S. **Utilização de polinômios segmentados na pesquisa zootécnica (considerações teóricas e práticas)**. Universidade Federal do Rio Grande do Sul – Faculdade de Agronomia – Departamento de Zootecnia, Seminário em Zootecnia, Porto Alegre-RS, 1989.
- SCHUMAKER, L.L. **Spline functions: basic theory**. Cambridge Mathematical Library, 3a edição, 2007.
- 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, 2011.
- SILVEIRA, F.G. **Classificação multivariada de modelos de crescimento para grupos genéticos de ovinos de corte**. Viçosa, UFV, 2010. 59p. Dissertação (Mestrado), Universidade Federal de Viçosa, 2011.
- SILVESTRE, A.M., PETIM-BATISTA, F., COLACO, J. The accuracy of seven mathematical functions in modeling dairy cattle lactation curves based on test-day records from varying sampling schemes. **Journal of Dairy Science**, v.89, p.1813–1821, 2006.
- SPEIDEL, S.E., ENNS, R.M., CREWS JR. D.H. Genetic analysis of longitudinal data in beef cattle: a review. **Genetic and Molecular Research**, v.9, n.1, p. 19-33, 2010.

- TORAL, F.L.B. **Avaliação do peso à desmama de bovinos cruzados Charolês-Zebu.** Viçosa, UFV, 2007. 130p. Tese (Doutorado) – Universidade Federal de Viçosa, 2007.
- TORRES JÚNIOR, R.A.A. **Markov chain Monte Carlo methods for estimating the covariance structure of longitudinal data - an application to dairy cattle data.** Cornell, CU, 2001. 91p. Tese (PhD) – Cornell University, 2001.
- VARGA, G. Teste de modelos estatísticos para a estrutura a termo no Brasil. **Revista Brasileira de Economia.** vol.63, n.4 p 361-394, 2009.
- WHITE, I.M.S., THOMPSON R., BROTHERSTONE S., Genetic and environmental smoothing of lactation curves with cubic splines. **Journal Dairy Science**, vol. 82 p.632–638, 1999.