

NADSON OLIVEIRA DE SOUZA

**AVALIAÇÃO GENÉTICA DA PRODUÇÃO E COMPOSIÇÃO DO LEITE EM
CAPRINOS DA RAÇA ALPINA UTILIZANDO MODELOS DE REGRESSÃO
ALEATÓRIA MULTICARACTERÍSTICOS**

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

VIÇOSA
MINAS GERAIS – BRASIL
2015

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

T

S725a
2015 Souza, Nadson Oliveira de, 1988-
Avaliação genética da produção e composição do leite em caprinos da raça alpina utilizando modelos de regressão aleatória multicaracterísticos / Nadson Oliveira de Souza. – Viçosa, MG, 2015.

xi, 60f. : il. (algumas color.) ; 29 cm.

Orientador: Robledo de Almeida Torres.
Dissertação (mestrado) - Universidade Federal de Viçosa.
Inclui bibliografia.

1. Leite de cabra. 2. Leite - Composição. 3. Leite - Produção. I. Universidade Federal de Viçosa. Departamento de Zootecnia. Programa de Pós-graduação em Zootecnia. II. Título.

CDD 22. ed. 637.17

NADSON OLIVEIRA DE SOUZA

**AVALIAÇÃO GENÉTICA DA PRODUÇÃO E COMPOSIÇÃO DO LEITE EM
CAPRINOS DA RAÇA ALPINA UTILIZANDO MODELOS DE REGRESSÃO
ALEATÓRIA MULTICARACTERÍSTICOS**

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

APROVADA: 23 de fevereiro de 2015.

Moyses Nascimento

Fabyano Fonseca e Silva
(Coorientador)

Robledo de Almeida Torres
(Orientador)

Aos meus pais Olivio e Sirlei,

“Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito. Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes”

Marthin Luther King

Dedico esse trabalho,

“Ninguém poderá jamais aperfeiçoar-se, se não tiver o mundo como mestre.
A experiência se adquire na prática”

William Shakespeare

AGRADECIMENTOS

À DEUS, pela vida, saúde e por todos os motivos que me levam a crer que sou a pessoa mais feliz e realizada desse mundo.

À Universidade Federal de Viçosa pela oportunidade da realização do Mestrado no Programa de Pós-Graduação em Zootecnia e ao Departamento de Zootecnia por fornecer a estrutura de execução da pesquisa

À FAPEMIG pela concessão da bolsa de estudos e financiamento de projetos correlacionados.

Ao meu orientador professor Dr. Robledo de Almeida Torres, que me recebeu muito bem no início do estágio e, desde então, é disponível para ensinamentos, conversas e conselhos nesse processo de formação pessoal e profissional.

Ao professor Dr. Marcelo Teixeira Rodrigues pela concessão dos dados do setor de Caprinocultura da Universidade Federal de Viçosa.

Ao Dr. Felipe Gomes da Silva, que teve a ideia do presente trabalho e, desde então, é disponível para ensinamentos e conversas. Um grande amigo que tenho desde o período de minha graduação até o presente momento.

Ao professor Dr. Fabyano Fonseca e Silva, pela orientação, amizade, risadas, paciência e dedicação nos ensinamentos estatístico e coorientação.

Ao Laboratório de Biometria, que forneceu a estrutura física para a execução das análises, e à FAPEMIG por proporcionar recurso financeiro para o estabelecimento da estrutura computacional.

Aos meus pais, Sirlei Cedras Oliveira de Souza e Olivio Ferreira de Souza, exemplo de amor e caráter. Muito obrigado por toda educação e esforço para que hoje fosse possível. Pelo incentivo e compreensão que eu precisava percorrer meu caminho para alcançar todos os meus objetivos. Apesar de estarem longe, são eles o motivo de eu seguir na luta cotidiana.

Às minhas irmãs, Naiana Oliveira e Naiara Augusta, pelo amor, carinho, pela torcida e por mesmo longe, estarem sempre presente ao longo de todo esse percurso.

Aos amigos, Vinicius Junqueira, André Mauric, Leonardo Peixoto, Leonardo Bhering, Humberto Fanelli e Giovani Caetano pelas horas de companhia e risadas nesse processo de formação em que todos estamos longe de casa.

Aos amigos e colegas de trabalho do Grupo de Melhoramento Genético de Caprinos (UFV), que me ajudaram na coleta dos dados e análises: Alessandra Silva, Ana Lúcia, Bárbara Félix, Delvan Silva, Geraldo Iria, Hinayah Rojas, Hugo Teixeira, Jô, Lais Brito, Luiz Brito, Mariana, Naila, Otávio, Paula, Yara, William Mariano, e demais que por descuido eu tenha esquecido. Foi muito bom conviver e aprender com vocês ao longo desses anos.

Aos meus amigos do melhoramento animal que de alguma forma me ajudou na realização deste trabalho, pelas trocas de conhecimento e ricas discussões: Edson Vinícius, Rodrigo Mota, Aline Cispim, Luciano Pinheiro, Jeferson Corrêa, Matilde Pessoa, Renata Veroneze, Sirlene Lázaro, Leonardo Gloria, Ederson Camargo, Rogério Velosos, Raphael Rocha, Lucas Verardo, labtec & cia.

Aos amigos de república (seca litro): Anderson Miranda, Carlos Junior, Cassio Machado e Fernando Gardioli, que sempre foram sinônimos de apoio, companheirismo, alegrias, convivência e troca de conhecimento.

Aos funcionários do setor de caprinocultura da UFV, principalmente Antônio, João, José Maria, Ronaldo, Anderson e aos demais, que foram muito importantes na execução da parte pratica deste trabalho.

Aos professores e funcionários do programa de pós-graduação em zootecnia por sempre terem me atendido muito bem e auxiliado nas duvidas e resolução de problemas.

A todos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

BIOGRAFIA

NADSON OLIVEIRA DE SOUZA, filho de Olivio Ferreira de Souza e Sirlei Cedras Oliveira de Souza, nasceu em 27 de abril de 1988, na cidade de Jacobina, Bahia.

Iniciou a graduação em Zootecnia na Universidade Federal do Vale do São Francisco (UNIVASF) no primeiro semestre de 2008. No final do primeiro semestre de 2010, transferiu para a Universidade Federal de Viçosa (UFV).

Durante o curso de graduação na Ufv se dedicou aos estudos e estagio na área de melhoramento animal.

No primeiro semestre de 2013 recebeu o título de Zootecnista pela Universidade Federal de Viçosa (UFV) e em abril do mesmo ano iniciou o curso de mestrado em Zootecnia na área de Melhoramento Animal pela na mesma instituição.

Em 23 de fevereiro de 2015, submeteu-se aos exames finais de defesa de dissertação para obtenção do título de *Magiste Scientiae* em Zootecnia.

SUMÁRIO

RESUMO	viii
ABSTRACT	x
INTRODUÇÃO GERAL	x
REFERÊNCIAS BIBLIOGRÁFICAS	6
CAPÍTULO I	8
Avaliação genética da produção e dos constituintes do leite de caprinos da raça Alpina utilizando modelos de regressão aleatória unicaracterístico.....	8
INTRODUÇÃO	10
MATERIAL E MÉTODOS	11
RESULTADO	15
DISCUSSÃO	37
CONCLUSÕES	40
REFERÊNCIA BIBLIOGRAFICA	41
CAPÍTULO II	44
Avaliação genética da produção e dos constituintes do leite de caprinos da raça Alpina utilizando modelos de regressão aleatória multicaracterísticos	44
INTRODUÇÃO	46
MATERIAL E MÉTODOS	47
RESULTADO	51
DISCUSSÃO	56
CONCLUSÕES	58
REFERÊNCIA BIBLIOGRAFICA	58
CONCLUSÕES GERAIS	60

RESUMO

SOUZA, Nadson Oliveira de, M.Sc., Universidade Federal de Viçosa, Fevereiro de 2015. **Avaliação da produção e composição do leite em caprinos da raça alpina utilizando modelos de regressão aleatória multicaracterístico.** Orientador: Robledo de Almeida Torres. Coorientador: Fabyano Fonseca e Silva.

O objetivo do presente estudo foi comparar modelos de diferentes ordens de ajuste por meio de funções polinomiais de Legendre, sob modelos de regressão aleatória unicaracterístico e multicaracterístico, com finalidade de obter modelos mais adequados para descrever as mudanças nas variâncias associadas a produção de leite, teor de proteína, gordura e lactose no dia do controle e estimar componentes de (co)variância e parâmetros genéticos. Foram utilizados 21.994 informações de produção de leite no dia do controle (pldc) de 774 cabras e para proteína, gordura e lactose foram utilizados 5.482, 6.353 e 5.500 registros de 588, 608 e 588, respectivamente, de cabras alpina de primeira lactação, em análise de regressão aleatória unicaracterístico (MRAU). Para análise de regressão aleatória multicaracterístico (MRAM) foram utilizados 34.635 informações de 644 cabras da raça alpina. Para as análises, foram incluídos os efeitos aleatório genético aditivo, ambiente permanente e residual. Além disso, os grupos contemporâneos (ano-estação, tipo de parto, agrupamento genético) e os efeitos linear e quadrático da idade da cabra ao parto foram incluídos como efeitos fixos. Os critérios utilizados para seleção dos modelos foram o logaritmo do máximo da função de verossimilhança (Log L), critério da informação de Akaike (AIC), critério da informação Bayesiano (BIC), teste da razão de verossimilhança (TRV) e número de parâmetros (NP). O programa Wombat foi utilizado em todas as análises genéticas. O modelo de regressão aleatória unicaracterístico utilizando polinômios ortogonais de Legendre de ordem cinco para a curva fixa, de ordem três para efeito genético aditivo, de ordem seis para o de ambiente permanente e ao menos cinco classes de variância residual foi o mais indicado para avaliação genética da produção de leite no dia do controle. Para teor de proteína o mais indicado foi o que considerou ordem cinco para curva fixa, ordem três para genético aditivo, ordem seis para ambiente permanente e três classes de variância residual. O

modelo que considerou a ordem cinco para a curva fixa, ordem seis para genético aditivo, ordem três para ambiente permanente e três classes de variância residual foi o mais indicado para a característica teor de gordura. Já para teor de lactose o modelo indicado para avaliação genética foi o que considerou a ordem três para curva fixa, ordem três para o genético aditivo, ordem seis para ambiente permanente e três classes de variância residual. Na análise utilizando modelo de regressão aleatória multicaracterístico o mais indicado para avaliação genética das características pldc, teor de proteína, gordura e lactose foi o que considerou a ordem três para a curva fixa, ordem três para efeito genético aditivo, ordem cinco para ambiente permanente e três classes de variância residual. Os componentes de (co)variância e os coeficientes de herdabilidade foram semelhantes em MRAU e MRAM ao longo da curva de lactação para as características pldc, proteína e lactose, já para gordura algumas diferenças foram observadas. Estudos que envolva o MRAM precisam ser mais frequentes, para que se possa entender o comportamento das características em análise conjunta e assim definir se vantagens existem em sua utilização em vez da utilização dos MRAU.

ABSTRACT

SOUZA, Nadson Oliveira de, M.Sc., Universidade Federal de Viçosa, February of 2015. **Evaluation of the production and composition of milk in the Alpine breed goats using multiple-trait random regression models.** Adviser: Robledo de Almeida Torres. Co-adviser: Fabyano Fonseca e Silva.

The study aimed was to compare different models of adjusting orders by Legendre polynomials and estimate variance components and genetic parameters, for single and multiple-trait random regression models for production milk, protein, fat and lactose. Were used 21,994 test-day (TD) records of milk from 774 goats, and protein, fat and lactose were used 5,482, 6,353 and 5,500 records from 588, 608 and 588, respectively, of Alpine goats in single trait random regression models (RRM). For multiple trait random regression models (RRMM) were used 34,635 records of 644 goats. For the analysis, the random effects were additive genetic, permanent environmental and residual. The fixed effects were contemporary groups (year-season, type of delivery, genetic group), the linear and quadratic effects of the covariable age of the goat at calving. We used the maximum of likelihood function, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Likelihood Ratio Test (LRT) and number of estimated parameters as criteria model choice. We used wombat program for all genetic analysis. The random regression model using polynomials Legendre of five order to the fixed curve, three order for additive genetic effect, six order for permanent environmental and five classes of residual variance was the most suitable for genetic evaluation of milk production test day. For protein the best was the model with five order for fixed curve, three order for additive genetic, six order for permanent environmental and three classes for residual variance. The model with five order for fixed curve, six order for additive genetic, four order for permanent environmental and three classes of residual variance was the most suitable for the fat content. For the lactose content the best model for genetic evaluation was with three order for fixed curve, three order for the additive genetic, six order for permanent environmental and three residual variance. In the analysis using random regression model multiple-trait the most suitable for genetic evaluation of TD

characteristics, protein, fat and lactose was deemed to three order for the fixed curve, order three for additive genetic effect, order five for permanent environmental and four classes residual variance. The variance components and heritability coefficients were similar in RRM and RRMM throughout the lactation curve for production milk characteristics, protein and lactose, while for fat there were some differences. Studies involving the RRMM need to be more frequent that one can understand traits the behavior in joint analysis and thus determine whether there are advantages in their use instead of using RRM.

INTRODUÇÃO GERAL

A caprinocultura leiteira brasileira ainda não é uma atividade de destaque comparado a bovinocultura leiteira, mas tem aumentado de forma bastante significativa sua participação no cenário agropecuário brasileiro, superando o constante desafio de conquistar e manter novos mercados para o leite e seus derivados.

Após a Instrução Normativa 37/2000 entrar em vigor, inúmeros estabelecimentos registrados nos Serviços de Inspeções passaram a produzir e comercializar leite pasteurizado, UHT, em pó e seus derivados como doces, iogurtes, queijos e etc.

Esta oferta cada vez mais variada de produtos tem exigido maior eficiência de todos aqueles envolvidos na atividade e, nesse sentido, devem ser considerados dois pontos de fundamental importância. O primeiro ponto é a qualidade e o segundo é a produtividade (Borges & Bresslau, 2002).

É de conhecimento que os teores dos constituintes do leite são atributos importantes para a indústria láctea, já que estão relacionados com o rendimento e a qualidade dos produtos. Embora existam poucos relatos de que preços melhores estejam sendo pagos pelo leite com maior teor de gordura ou de proteína, espera-se que no futuro, tais características passem a ser beneficiadas e com a determinação de valores econômicos possam ser criados índices de seleção para o programa de melhoramento de caprinos. Desta forma, é importante que os melhoristas estejam à frente desta tendência, selecionando os melhores modelos de avaliação e estimando parâmetros genéticos destas características (Silva, 2011).

Modelos de Regressão Aleatória

Em caprinos de leite, as características de produção e qualidade do leite podem ser mensuradas em um mesmo animal repetidas vezes ao longo da sua vida e atualmente essa mensuração vem sendo realizada utilizando as produções no dia do controle (pldc), o que possibilita uma avaliação genética

mais acurada por considerar efeitos ambientais específicos (efeitos não genéticos) que podem estar influenciando o desempenho dos animais nos dias do controle (Jamrozik et al., 1997; Strabel & Jamrozik, 2002). A avaliação genética para produção e qualidade do leite tem sido realizada de diversas formas, sendo encontrados normalmente trabalhos utilizando modelos de repetibilidade, modelos multicaracterísticos e, mais recentemente, modelos de regressão aleatória.

O modelo de Repetibilidade é a forma mais simples de análise de dados longitudinais, ele assume que todas as medidas ao longo do tempo são a mesma característica. Esse modelo pressupõe que as variâncias genéticas e não genéticas, não variam ao longo do tempo, portanto, todas as correlações entre as medias repetidas sejam iguais a unidade, ou seja, assume-se que os mesmos genes controlam o desempenho ao longo do tempo. Faro (2003), ao estudar produção de leite em bovinos da raça Caracu, verificou que modelo de repetibilidade não seria adequado para a avaliação genética.

Outra forma de analisar características longitudinais seria por meio de análises multicaracterísticos. Nesses modelos, assume-se que cada mensuração feita em várias idades do indivíduo é tratada como característica distintas, elevando o número de parâmetros a serem estimados, e que as covariâncias entre as medições, nas diferentes idades, variam e as correlações podem ser diferentes da unidade; no entanto, nenhuma pressuposição é feita sobre a estrutura de covariâncias (Samento, 2007). Nesses modelos, consideram-se associações entre as características e utilizam-se correlações genéticas e fenotípicas na avaliação da característica (Silva et al., 2008).

Os modelos de regressão aleatória (MRA), metodologia proposta inicialmente por Henderson Jr. (1982), têm sido propostos como uma alternativa para modelar características longitudinais. Neste modelo, o que difere dos tradicionais modelos (repetibilidade e multicaracterístico), está na inclusão de colunas de covariâncias na matriz de incidência dos efeitos aleatórios, em substituição às colunas de zero e um. Os MRA consideram as mudanças ocorridas ao longo da vida do animal e permitem a estimação dos componentes de (co)variância e conseqüente estimação dos parâmetros

genéticos em qualquer momento da lactação, além de identificar a fase em que há maior variabilidade genética.

Além disso, os MRA usufruem de todas as vantagens atribuídas ao uso de dados longitudinais. Pode citar, como aspectos positivos do MRA: não exige número mínimo de medidas por animal; considera o parentesco existente entre os animais; as análises abrangem todo o intervalo contínuo em que as medidas foram tomadas, não permitindo extrapolação; permitem prever o valor genético do animal para as características em qualquer ponto desse intervalo; possibilitam obter estimativas de (co)variâncias para as características, nas quais não há observações por meio de interpolação (Menezes, 2010).

O uso dos MRA requer a escolha e utilização de uma função que descreva a ação dos efeitos fixos e aleatórios que atuam sobre as características estudadas. Dentre os polinômios ortogonais e as funções paramétricas, os polinômios de Legendre e a função de Wilmink têm sido os mais utilizados para modelar a curva de lactação. Segundo Brotherstone (2000), as funções paramétricas descrevem melhor a curva de lactação, entretanto resultaram em correlações genéticas negativas entre produções no período inicial e final da lactação. Quando se ajustaram os polinômios de Legendre, essas correlações foram positivas ao longo de toda a lactação. Schaeffer (2004) relata que os polinômios ortogonais são os mais apropriados para serem usados, como covariáveis, pelos modelos de regressão aleatória.

Os polinômios ortogonais de Legendre (POL) são funções não paramétricas amplamente utilizadas para estudar a estrutura de covariância para efeitos aleatório genéticos aditivos e de ambiente permanente, e estes modelos não requerem qualquer pressuposição sobre a forma da curva ou trajetória da característica com o tempo.

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 POL 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.

Modelos de Regressão Aleatória Multicaracterísticos

Muitos estudos têm sido feitos com o objetivo de estimar componentes de (co)variância e parâmetros genéticos para produção e componentes do leite no dia do controle obtidos por modelos de regressão aleatória em bovinos de leite e búfalos (Aspilcueta-Borquis et al., 2010), mas poucos estudos aplicam modelos de regressão aleatória multicaracterístico (MRAM) para produção de leite, proteína, gordura e lactose no dia do controle. A implementação de análises envolvendo um MRAM é interessante, pois permite o estudo das relações genéticas entre a produção e os componentes do leite ao longo da lactação.

Jamrozik et al. (1997), testaram a utilização de Modelos de Regressão Aleatória Multicaracterísticos (MRAM) para analisar a PLDC, a produção de gordura (PG), a produção de proteína (PP) e a contagem de células somáticas (CCS), dentro de três lactações. Os autores chamaram esta tentativa de escolha lógica de teste frente às mudanças do sistema de coletas de dados produtivos do rebanho canadense. Espera-se dos MRAM melhores parâmetros genéticos com o aumento de características contínuas de interesse econômico a serem avaliadas em um rebanho.

Na conclusão apresentada por Jamrozik e seus colaboradores é possível perceber que a utilização de MRAM é muito interessante, pois permite aproveitar melhor diversas características longitudinais com grande volume de dados perdidos. Dessa forma se obtêm predições muito mais acuradas em relação ao modelo de repetibilidade, para os valores genéticos destas características.

Crítérios para escolha de um modelo

Na escolha do melhor modelo de regressão aleatória, contrasta-se o uso de diferentes equações para modelar a curva de produção de leite, e diferentes

classes de variâncias residuais. Entretanto, critérios são necessários para que o melhor modelo seja escolhido para a situação analisada.

Portanto, a comparação de modelos e a verificação de suas eficácias não é uma tarefa tão fácil, necessitando seguir critérios para que escolhas arbitrárias não sejam feitas. São encontrados na literatura alguns critérios para contrastar MRA, que são: critérios do logaritmo do máximo da função de verossimilhança (Log L), critério da informação de Akaike (AIC), critério da informação Bayesiano de Schwarz (BIC) e teste da razão da verossimilhança (TRV).

Com o intuito de escolher modelos menos parametrizados, maior facilidade de convergência, menor tempo de análise, menor erro nas estimativas dos parâmetros e menores erros de predições causada por ineficiência do conjunto software-hardwares, outro critério também utilizado é o número de parâmetros (NP).

O AIC é um critério de informação proposto por Hirotugu Akaike em 1971 com o objetivo de auxiliar a seleção de modelos estatísticos. A teoria está descrita em Akaike (1973), em que o autor demonstra que a adição de $2p$, onde p representa o número de parâmetros no modelo, pode corrigir o estimador do máximo logaritmo de verossimilhança ($-2 \times \text{Log } L$), o que é interessante também do ponto de vista de penalizar a parametrização excessiva do modelo. O AIC não pode ser considerado um teste, por não conter regras de valores críticos. Sua utilidade está em comparar modelos, sendo calculado um valor de AIC para cada modelo avaliado, dentre esses, o que apresentarem menores valores serão melhores. Sua fórmula é:

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

em que:

L é o ponto do máximo da função de verossimilhança; p corresponde ao número de parâmetros no modelo.

O BIC, critério de informação proposto por Gideon E. Schwarz em 1978, é considerado semelhante ao AIC, entretanto, geralmente penaliza mais a adição de novos parâmetros. Modelos que apresentam menores valores serão

melhores. Sua teoria está demonstrada em Schwarz (1978). A seguir está exposta a fórmula para se calcular o BIC.

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

em que:

p é o número de parâmetros estimados; N é o número de observação;

r é o posto da matriz X (matriz de incidência dos efeitos fixos);

$\text{Log } L$ é o máximo logaritmo da função de verossimilhança.

O teste da razão de verossimilhança (TRV) compara as diferenças entre o máximo das funções de verossimilhança maximizadas de cada modelo e baseia-se na razão entre o logaritmo neperiano do máximo de duas funções de verossimilhança, exemplo: $\ln L_1$ e $\ln L_2$. Então a estatística do teste será:

$$TRV = -2 [\ln L_1 - \ln L_2] \sim X_g^2$$

em que:

g é o número de graus de liberdade com base no $NP_{\text{completo}} - NP_{\text{reduzido}}$

O teste é distribuído assintoticamente e o valor crítico é obtido da tabela qui-quadrado (X^2) com g grau de liberdade, se o valor calculado for maior que o tabelado o teste é significativo e o modelo completo fornece melhor ajuste, em relação ao modelo reduzido. A hipótese de nulidade testada foi de que os modelos, completo e reduzido, não diferiam entre si.

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.
- ASPILCUETA-BORQUIS, R. B., R. C. SESANA, M. MUNOZ-BERROCAL, L. O. SENO, A. B. BIGNARDI, L. EL FARO, L. G. ALBUQUERQUE, G. M. CAMARGO, AND H. TONHATI. 2010. Genetic parameters for milk, fat and protein yields in Murrah buffaloes (*Bubalus bubalis* Artiodactyla, Bovidae). **Genet. Mol. Biol.** 33:71–77.

- BORGES, C.H.P.; & BRESSLAU, S. **Produção de leite de cabra em confinamento**. VI Simpósio de Pecuária do Nordeste e III Semana da Caprino-ovicultura Brasileira. Fortaleza-CE, 2002.
- COBUCI, J.A.; COSTA, C.N.; TEIXEIRA, N. et al. **Estimativas de herdabilidade para produção de leite no dia do controle de vacas da raça Holandesa utilizando regressão aleatória**. In: V SIMPÓSIO DA SOCIEDADE BRASILEIRA DE MELHORAMENTO ANIMAL, 2004, Pirassununga. **Anais...** Pirassununga: Sociedade Brasileira de Melhoramento Animal, 2004.
- HENDERSON Jr., C.R. Analysis of covariance in the mixed model: higher-level, nonhomogeneous, and random regressions. **Biometrics**, Washington, v.38, p.623-640, 1982.
- JAMROZIK, J.; KISTEMAKER, G.J.; DEKKERS, J.C.M. et al. Comparison of possible covariates for use in a random regression model for analyses of test day yields. **Journal of Dairy Science**, v.80, n.10, p.2550-2556, 1997.
- LIU, Y.X.; ZHANG, J.; SCHAEFFER L.R. et al. Short communication: Optimal random regression models for milk production in diary cattle. **Journal of Dairy Science**, Champaign, v.89, p.2233-2235, 2006.
- 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.
- SAMENTO, J. L. R. **Modelos de regressão aleatória para avaliação genética da curva de crescimento de ovinos da raça Santa Inês**. Viçosa, UFV, 2007. 101p. Tese (Doutorado) – Universidade Federal de Viçosa, 2007.
- SCHWARZ, G. Estimating the Dimension of a Model. **Annals of Statistics**, 1978, 6, p.461-464.
- 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.
- SILVA, M. A.; THIÉBAUT, J. T. L.; VALENTE, B. D. et al. **Modelos lineares aplicados ao melhoramento genético animal**. FEPMVZ Editora. Escola de Veterinária da UFMG. 375p. 2008.
- STRABEL, T.; JAMROZIK, J. The effect on incorrect estimated variance covariance components on genetic evaluation of dairy cattle with random regression models. In: WORLD CONGRESS GENETIC APPLIED LIVESTOCK PRODUCTION, 7, 2002, Montpellier, France. **Proceeding**. Montpellier, CD-ROM, 2002.

CAPÍTULO I

Avaliação genética da produção e dos constituintes do leite de caprinos da raça Alpina utilizando modelos de regressão aleatória unicaracterístico

RESUMO: Os parâmetros genéticos para características produtivas em caprinos leiteiros são estimados em função da produção do leite e seus componentes com base na produção de leite no dia do controle (pldc), com a utilização de modelos de regressão aleatória (MRA). Diante disso, o objetivo do presente estudo foi comparar MRA com diferentes ordens de ajuste por meio de funções polinomiais de Legendre com a finalidade de selecionar o modelo mais adequado, estimar os componentes de (co)variância e os parâmetros genético para produção de leite, teor de proteína, gordura e lactose em cabras da raça alpina. Foram utilizados 21.994 informações de pldc de 774 cabras, já para proteína, gordura e lactose foram utilizados 5.482, 6.353 e 5.500 registros de 588, 608 e 588 cabras, respectivamente. Para a análise genética e fornecimento dos componentes de (co)variância foi utilizado o programa Wombat. Os critérios logaritmo do máximo da função de verossimilhança (Log L), critério da informação de Akaike (AIC), critério da informação Bayesiano (BIC), teste da razão de verossimilhança (TRV) e número de parâmetros (NP) foram utilizados para escolha do melhor modelo. Para pldc, o modelo que considerou ordem 5, 3 e 6 para curva fixa e os efeitos aleatório genético aditivo e ambiente permanente, respectivamente, e 5 classes de variância residual foi o mais adequado. Já para proteína o modelo de ordem 5, 3, 6 para curva fixa, para o genético aditivo e para o ambiente permanente, e três classes de variância residual foi o que melhor se ajustou aos dados. Para gordura o modelo indicado foi o de ordem 5, 6 e 3 para curva fixa e para os efeitos aleatórios, genético aditivo e ambiente permanente, respectivamente, e 3 classes de variância residual. Para lactose o mais indicado foi o que considerou a ordem 3, 3 e 6 para efeito fixo e para os efeitos aleatórios, genético aditivo e de ambiente permanente, respectivamente, e 3 classes de variância residual.

Palavras-chave: AIC, BIC, componente de (co)variância, herdabilidade

Genetic evaluation of test-day milk, protein, fat and lactose yield from Alpine goats using random regression models

ABSTRACT: Genetic parameters for production traits in dairy goats are estimated based on milk, protein, fat and lactose yield using test-day (TD) models by random regression models (RRM). This study aimed to identify the RRM of better fit adjusted by Legendre Polynomials with different orders, estimate variance components and genetic parameters of milk, protein, fat and lactose yield of dairy Alpine goats. We analyzed 21,994 TD records of milk, 5,482 of protein, 6,353 of fat and 5,500 of lactose from 774, 588, 608 and 588 goats, respectively. We used Wombat software to estimate genetic parameters and variance components. We used the maximum of likelihood function, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Likelihood Ratio Test (LRT) and number of estimated parameters as criteria model choice. The model that showed best fit for milk yield adjusted Legendre polynomials of 5, 3 and 6 orders for fixed effect, additive genetic and permanent environment random effects, respectively, and 5 classes of residual variance. The best model for protein yield was adjusted with the same Legendre polynomials orders for fixed effect and additive genetic and permanent environment random effects as those fitted for milk yield. However, three classes of residual effect presented better fit. The best model for fat was fitted with order 5, 6 and 4 for fixed, additive genetic and permanent environment effects, respectively, and three classes for residual variance. The model that showed best fit for lactose presented Legendre polynomials of order 3, 3 and 6 for fixed and additive genetic and permanent environment random effects, respectively, and three classes of residual variance.

Keywords: Akaike criteria information, Bayesian criteria information, variance components, heritability.

INTRODUÇÃO

A caprinocultura leiteira ainda não se tornou uma atividade que mereça destaque no cenário brasileiro no que diz respeito a produção de leite, mas nos últimos anos o setor vem crescendo por ser uma atividade potencialmente rentável, já que os caprinos são mais eficientes como produtores de leite do que de carne. Diante deste crescimento, a busca de animais superiores para quantidade e qualidade do leite têm sido objetivo de diversas pesquisas.

Atualmente, os parâmetros genéticos para características produtivas em caprinos leiteiros são estimados em função da produção de leite e seus componentes com base na produção de leite no dia do controle (pldc). Adicionalmente, com a entrada em vigor da Instrução Normativa 37, passou a ser exigido pelo menos uma análise mensal do leite, para averiguação da qualidade do mesmo com relação aos padrões mínimos de composição, dando origem então, a maior quantidade de informação dos componentes do leite ao longo da lactação.

Com a disponibilidade de dados durante toda lactação da cabra, tem sido sugerido a utilização de modelos de regressão aleatória (MRA) associados à metodologia de modelos mistos, adequada para avaliação de dados longitudinais, possibilitando a seleção de caprinos de aptidão leiteira pela pldc ao invés da produção de leite acumulada (PLA), metodologia esta proposta inicialmente por Henderson Jr. (1982). O uso dos MRA requer a escolha e utilização de uma função que descreva a ação dos efeitos fixos e aleatórios que atuam sobre as características estudadas, dentre estas funções, os polinômios ortogonais de Legendre têm sido eleitos (Liu et al., 2006).

Estudos realizados no Brasil relataram que entre esses polinômios os de ordem quatro e cinco têm sido os preferidos para modelar os efeitos genéticos e não genéticos que atuam na produção de leite em bovino (Costa et al., 2005; Cobuci et al., 2006 e Araújo et al., 2006). Já para caprinos de aptidão leiteira polinômios de ordem quatro e seis têm sido os preferidos para modelar os efeitos genéticos e não genéticos (Menezes et al., 2008 e Sarmiento et al., 2008).

Zampar (2012) analisou 27.988 informações de teores de proteína e gordura do leite de vacas de primeira lactação com o objetivo de identificar o melhor MRA utilizando polinômios ortogonais de Legendre. Dentre os modelos estudados, os de quinta e sexta ordem foram os que melhores se adequaram aos teores de gordura e o de quarta ordem para teor de proteína. Mais estudos para proteína, gordura e lactose do leite ainda necessitam serem realizados, já que ainda são escassos nas literaturas.

Diante do exposto, o objetivo do presente estudo foi comparar modelos de regressão aleatória com diferentes ordens de ajuste por meio de funções polinomiais de Legendre com a finalidade de escolher o modelo mais adequado e estimar parâmetros genéticos para produção de leite, de proteína, de gordura e de lactose, no dia do controle de cabras da raça Alpina.

MATERIAL E MÉTODOS

Os registros de produção e constituintes do leite de caprinos da raça Alpina utilizados neste estudo foram provenientes do setor de caprinocultura do Departamento de Zootecnia (DZO) da Universidade Federal de Viçosa (UFV) em Viçosa, Minas Gerais. A produção de leite no dia do controle (pldc) que representa a soma do leite coletado no período de 24 horas foi obtida por meio de controles leiteiros realizados semanalmente em ordenha mecânica, sendo duas vezes ao dia. A primeira coleta teve início às seis horas da manhã e a segunda coleta teve início às quatorze horas. Após o término do controle leiteiro de cada semana, os dados eram conferidos e digitados no banco de dados.

Os constituintes do leite utilizado foram obtidos a partir de amostras coletadas durante o período de 1999 à 2014, sendo realizada uma vez ao mês em recipiente de aproximadamente 75 ml, em que no período da manhã eram coletados a fração de 2/3 e a tarde coletados 1/3 do volume total, feito isso, os recipientes recebiam identificações das cabras e armazenados com conservantes e posteriormente enviados ao laboratório de qualidade do leite do DZO, que retornava os dados de gordura, proteína e lactose, além da

contagem de células somáticas. Estes dados eram conferidos e digitados no banco de dados.

A organização e preparação do arquivo de dados foram realizadas utilizando-se o software SAS 9.1. Após análise crítica dos dados, restaram para pldc 21.994 informações de 774 cabras, já para o arquivo de proteína, gordura e lactose restaram 5.482, 6.353 e 5.500 registros de 588, 608 e 588, respectivamente, de cabras da raça Alpina de primeira lactação. As lactações foram truncadas aos 290 dias e animais que tiveram o primeiro parto entre os 7 e 36 meses de idade foram mantidas.

Para corrigir erros de pedigree e recodificar os dados alfanuméricos foi utilizado o programa RENPED (Silva, 2011). Já o programa WOMBAT (Meyer, 2007), foi utilizado em todas as análises genéticas, fornecendo as estimativas dos parâmetros para cada animal avaliado e bem como os valores necessários para o cálculo dos critérios de escolha dos modelos.

As análises foram realizadas separadamente para pldc, proteína, gordura e lactose. No arquivo, apenas foram utilizados animais com composição genética (Alpina = 100% = 1; 90% ≤ Alpina < 100% = 2; 80% ≤ Alpina < 90% = 3; 70% ≤ Alpina < 80% = 4; 60% ≤ Alpina < 70% = 5; 50% ≤ Alpina < 60% = 6). O ano-estação do controle (de 2000 a 2014 e estação 1 de março a setembro e 2 para os demais meses), tipo de parto (até três filhotes) e idade da cabra ao parto como covariável, considerando-se os efeitos linear e quadrático.

O modelo utilizado para testar a influência dos fatores sobre as características de estudo foi:

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

em que:

Y_{ijkl} é o valor observado para a característica estudada na cabra l , tipo de parto i , grupo genético j , ano-estação k ;

μ é a média populacional da característica;

T_i é o efeito fixo do tipo de parto i ;

G_j é o efeito fixo do agrupamento genético j;

A_k é o efeito fixo do ano-estação k;

b_1 e b_2 é o coeficientes linear e quadrático de regressão da característica analisada em função da idade da cabra ao parto (ICP);

P_{ijkl} é a idade da cabra l ao parto;

\bar{P} é a média amostral da ICP;

ε_{ijkl} é o erro aleatório;

As análises estatísticas foram realizadas empregando-se o procedimento Generalized Linear Model (Proc GLM) do programa Statistics Analysis System (SAS, 2003). Para as características proteína e gordura, o efeito fixo T não foi incluída no modelo por não ser significativo ($P > 0,05$).

Os modelos testados continham combinações de distintas ordens de ajustes para os polinômios das curvas, fixa (F_), genética aditiva (A_) e de ambiente permanente (P_), além do número de classes de variâncias residuais (E_). Por exemplo, um modelo com regressão fixa de ordem 3, regressão genética aditiva de ordem 4 e regressão de ambiente permanente de ordem 5, além de considerar a variância residual heterogênea contendo 3 classes, pode ser resumido pelo código F3A4P5H3. 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 característica em estudo no controle j da cabra i; EF_i consiste no conjunto de efeitos fixos, constituído por grupamento genético, tipo de parto e ano-estação. A idade da cabra ao parto foi incluída como covariável, com efeito linear e quadrático; b_m é o coeficiente de regressão m da característica estudada 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 e K_γ são as ordens dos polinômios ortogonais de Legendre; t_{ij} é a variável

controle da cabra i padronizada para o intervalo -1 a 1 conforme descrito por Kirkpatrick et al. (1990); $\phi_m(t_{ij})$ é a função polinomial de Legendre para o parâmetro m avaliado para a idade t_{ij} ; e ε_{ij} consiste nos efeitos específicos à cada observação não explicados pelos fatores de correção nem pelas regressões do modelo.

Para a escolha dos modelos que melhores se ajustaram aos dados, os seguintes critérios foram utilizados seguindo a seguinte ordem: teste da razão de verossimilhança (TRV), critério da informação de Akaike (AIC) (Akaike, 1973) e critério da informação Bayesiano (BIC) (Schwarz, 1978), número de parâmetros (NP) e logaritmo do máximo da função de verossimilhança (Log L)

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 melhores ajustes do modelo. Os critérios de informação podem ser representados como: $AIC = -2 \log L + 2p$ e $BIC = -2 \log L + p \log(n - r)$,

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 Log L é o logaritmo do máximo da função de verossimilhança.

Vale ressaltar que os AICs e BICs que serão apresentados neste trabalho estão modificados (AIC_m e BIC_m) através da multiplicação pelo termo (-1/2), pois isso faz com que estes critérios fiquem em escala semelhante ao Log L e com que o maior valor corresponda ao melhor modelo, o que facilita a comparação dos modelos e comportamentos dos critérios frente a alterações nos modelos.

Para comparação entre modelos completos e reduzidos também foi utilizado o Teste da Razão de Verossimilhança (TRV) (RAO, 1973), ao nível de 1% de probabilidade. $TRV = -2 [\ln L_1 - \ln L_2]$

O TRV compara as diferenças entre o máximo das funções de verossimilhança linearizadas de cada modelo, com o valor obtido de uma

distribuição qui-quadrado com o número de grau de liberdade igual à diferença no número de parâmetros entre os modelos.

O número de parâmetros é um critério que foi utilizado na escolha do modelo, haja vista que modelos menos parametrizados são preferíveis devido ao menor tempo de análise e maior facilidade de convergência.

Quanto a utilização das estruturas de variâncias residuais, avaliou-se duas formas de estruturas, uma com variância residual constante ao longo da lactação e a segunda variando em classes de dias em lactação. Quando heterogênea, foram consideradas as seguintes classes: duas (do 5° ao 134° e do 135° aos 290° dias de lactação), três (do 5° ao 55°, do 56° ao 132° e do 133° aos 290° dias em lactação), quatro (do 5° ao 62°, do 63° ao 132°, do 133° ao 223° e do 224° aos 290° dias em lactação) e cinco (do 5° ao 42°, do 43° ao 71°, do 72° ao 105°, do 106° ao 189° e do 190° ao 290° dia em lactação).

RESULTADO

Critérios de avaliação de modelos

Na Figura 1 à 12 estão os resultados do Log L, AICm e BICm para as características gordura, proteína, lactose e pldc em função do número de classes de variância residual e das ordens dos polinômios para os efeitos genéticos aditivo, ambiente permanente e dos graus para a curva fixa. É possível perceber alterações nos valores do Log L, AICm e BICm com o aumento das ordens das regressões dos efeitos genéticos e de ambiente permanente, sendo que estas alterações possibilitaram incrementos cada vez menores. Enquanto que na curva fixa e no número de classes de variância residuais os incrementos foram em menores escala.

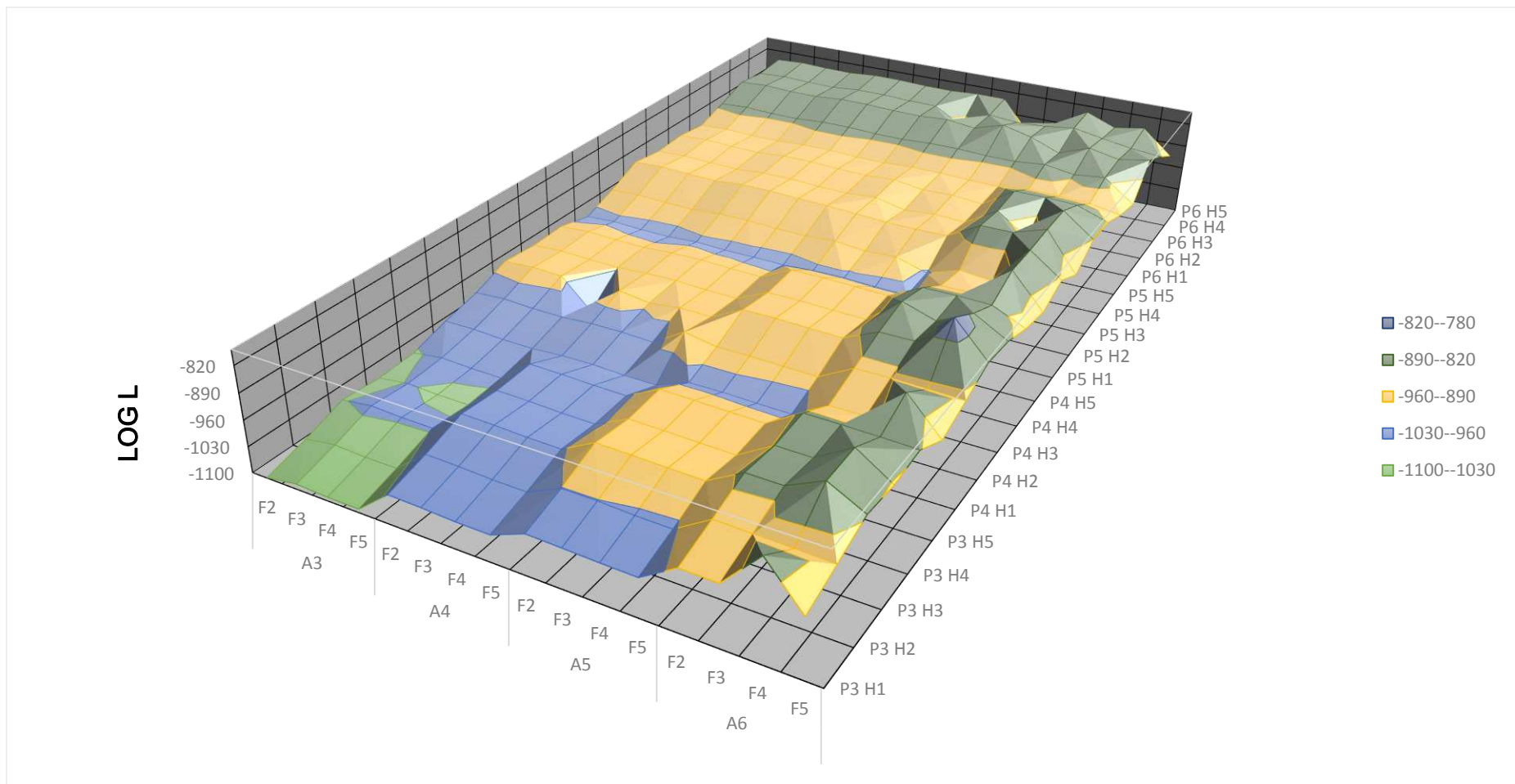


Figura 1: Resultados do logaritmo da função de máxima verossimilhança em função das ordens dos polinômios para os efeitos genéticos aditivo (A), ambiente permanente (P) e dos graus para a regressão fixa (F) e do número de classes de variância residual (H) para característica gordura.

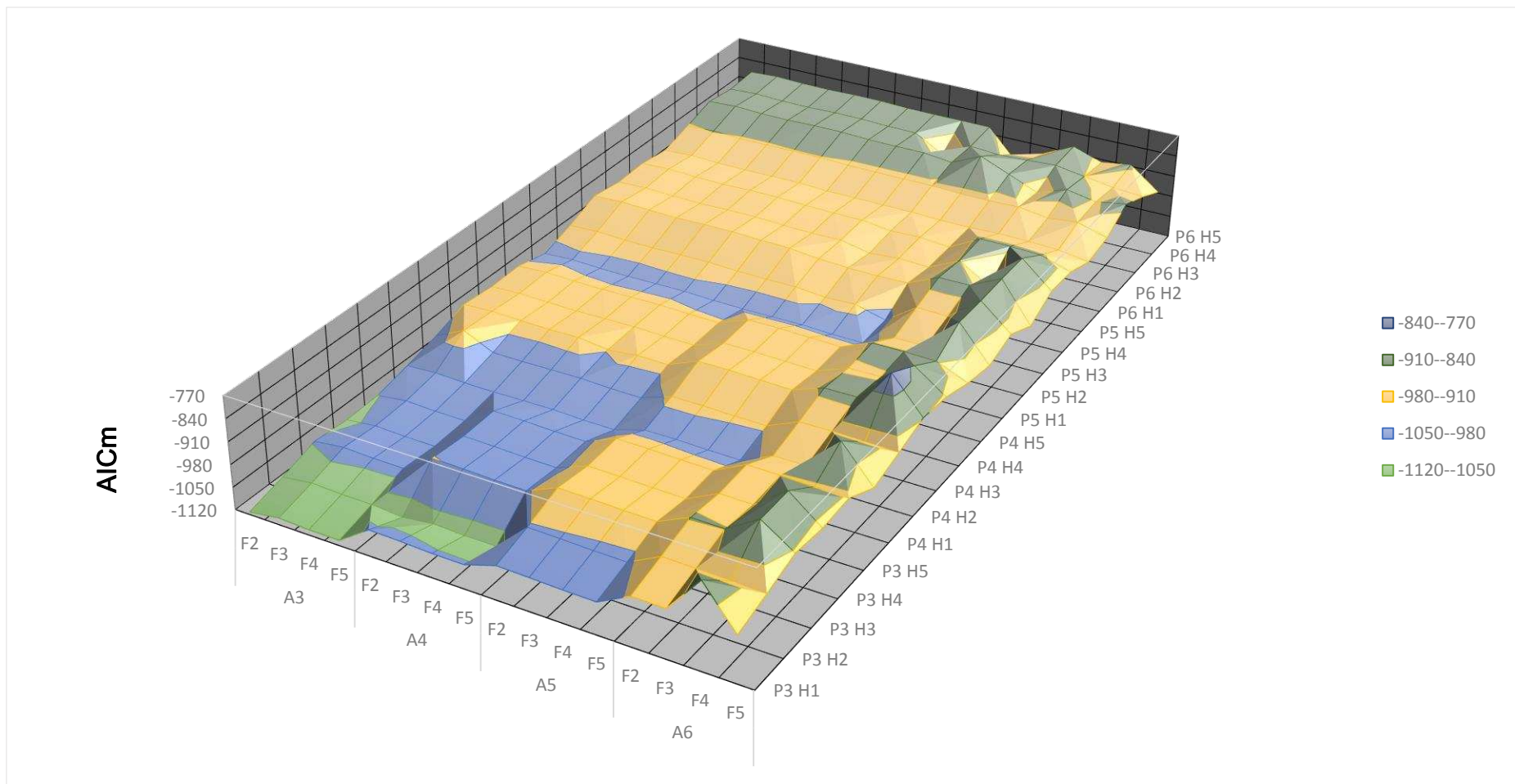


Figura 2: Resultados do critério da informação de Akaike modificado ($AIC_m = -1/2 AIC$) em função das ordens dos polinômios para os efeitos genéticos aditivo (A), ambiente permanente (P) e dos graus para a regressão fixa (F) e do número de classes de variância residual (H) para a característica gordura.

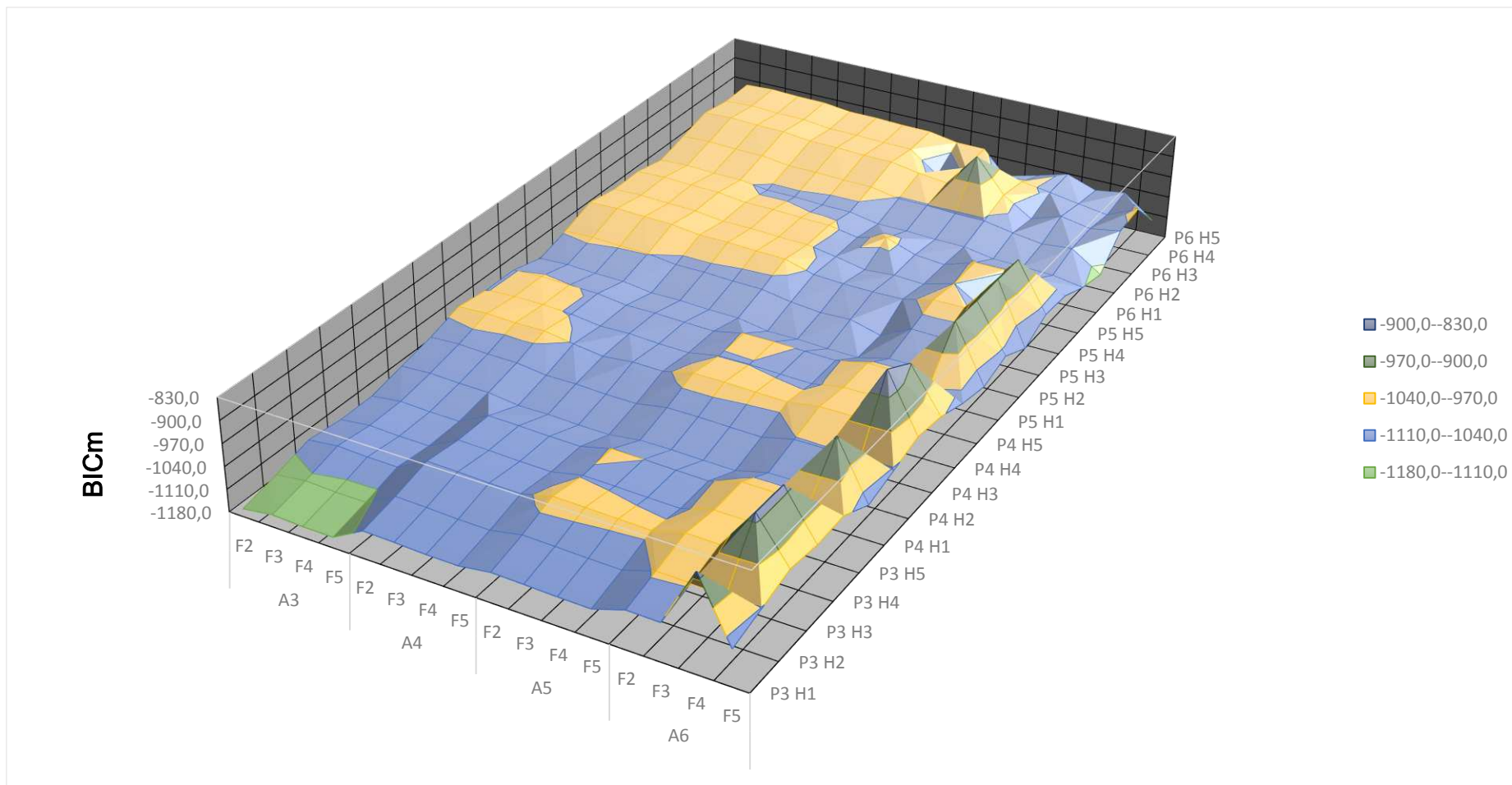


Figura 3: Resultados do critério da informação Bayesiano modificado ($BICm = -1/2BIC$) em função das ordens dos polinômios para os efeitos genéticos aditivo (A), ambiente permanente (P) e dos graus para a regressão fixa (F) e do número de classes de variância residual (H) para característica gordura.

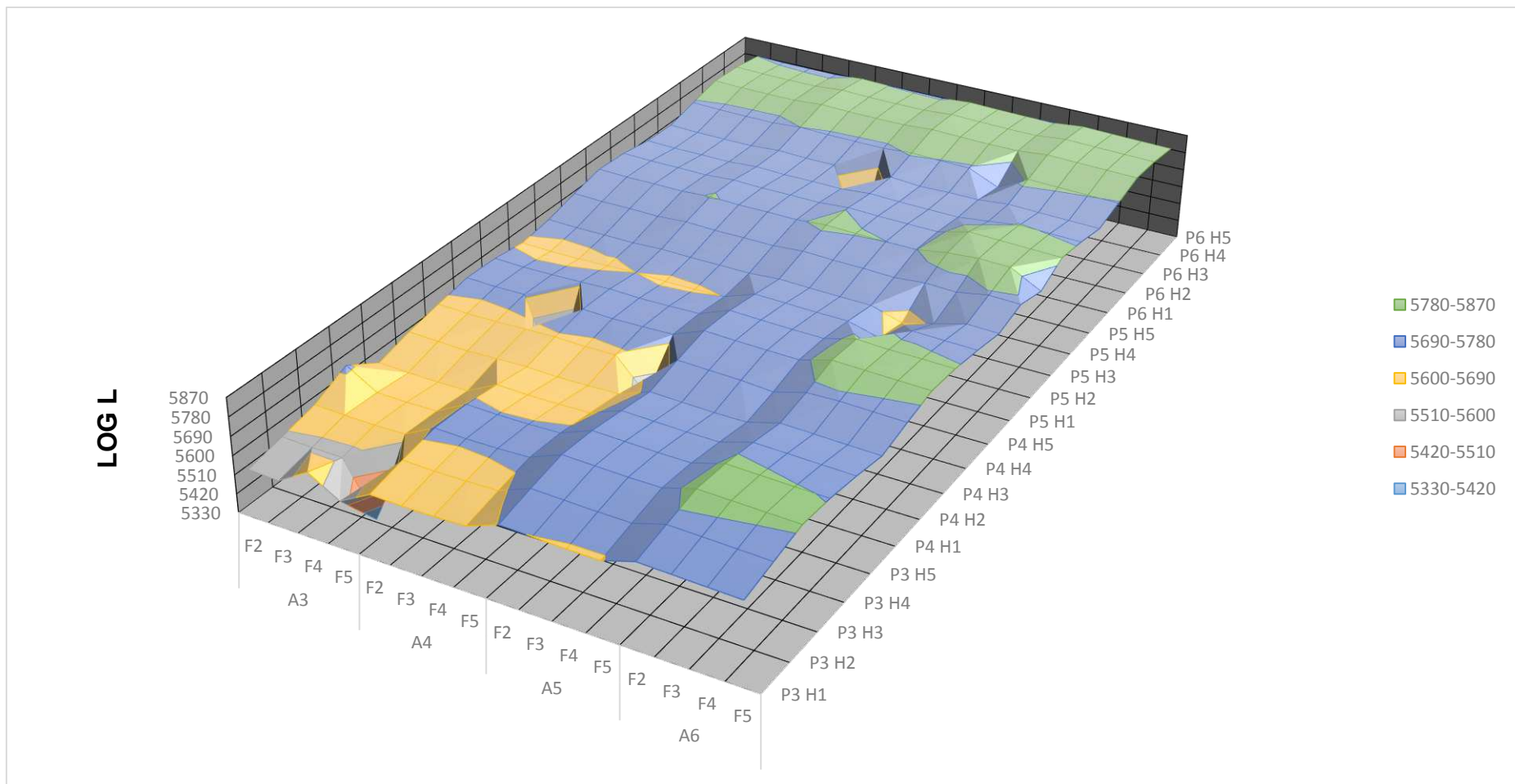


Figura 4: Resultados do logaritmo da função de máxima verossimilhança em função das ordens dos polinômios para os efeitos genéticos aditivo (A), ambiente permanente (P) e dos graus para a regressão fixa (F) e do número de classes de variância residual (H) para característica lactose.

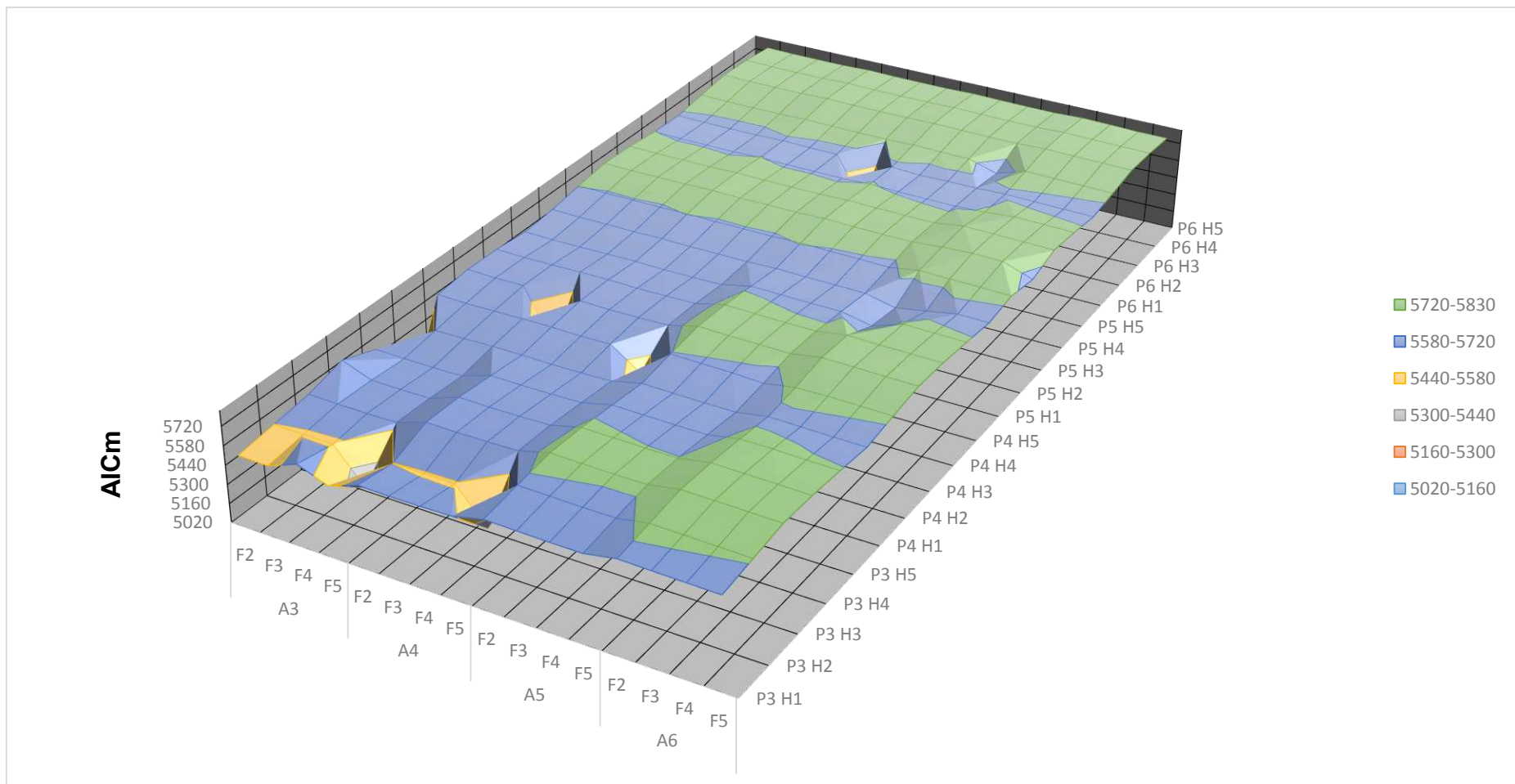


Figura 5: Resultados do critério da informação de Akaike modificado ($AIC_m = -1/2AIC$) em função das ordens dos polinômios para os efeitos genéticos aditivo (A), ambiente permanente (P) e dos graus para a regressão fixa (F) e do número de classes de variância residual (H) para a característica lactose.

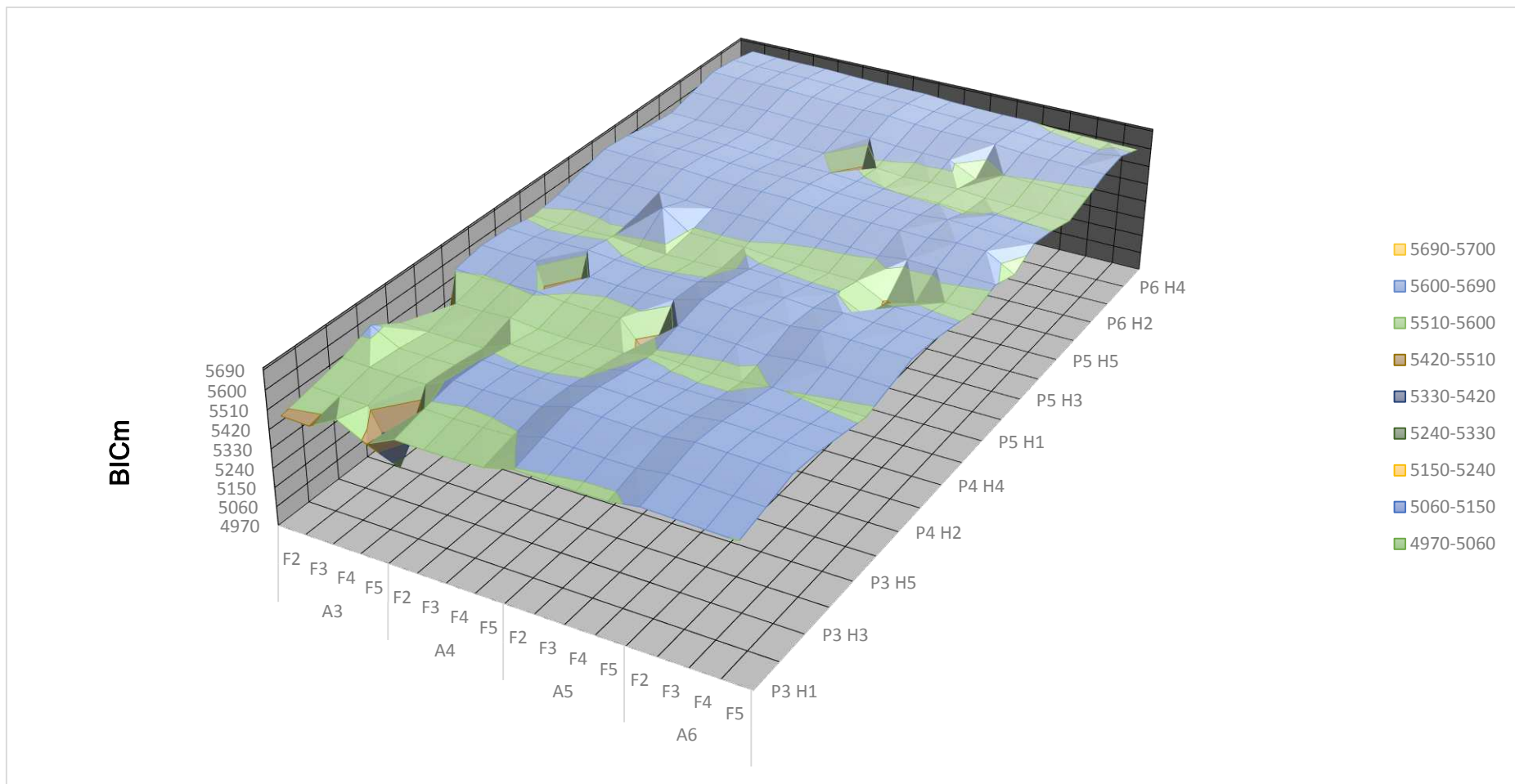


Figura 6: Resultados do critério da informação Bayesiano modificado ($BICm = -1/2BIC$) em função das ordens dos polinômios para os efeitos genéticos aditivo (A), ambiente permanente (P) e dos graus para a regressão fixa (F) e do número de classes de variância residual (H) para característica lactose.

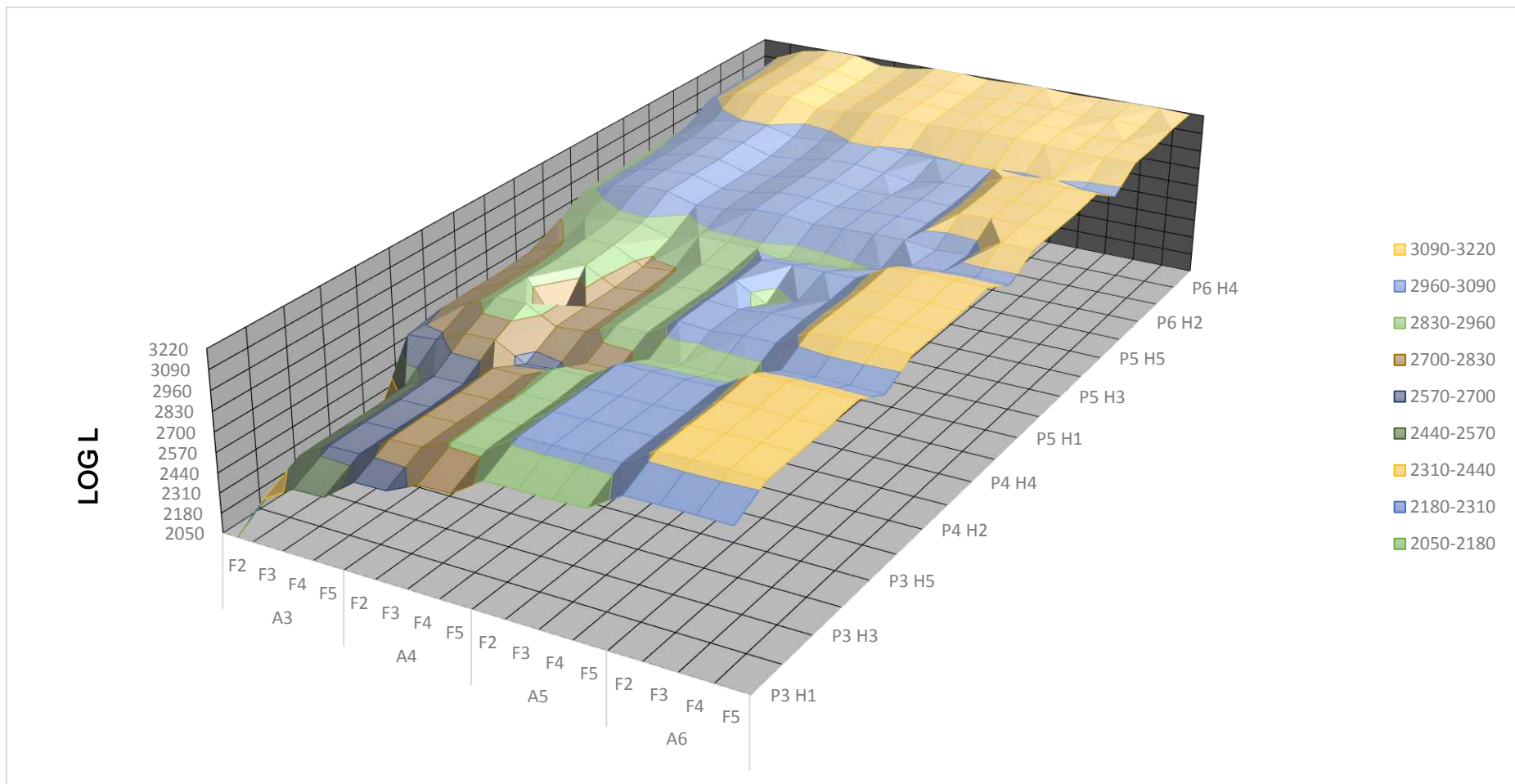


Figura 7: Resultados do logaritmo da função de máxima verossimilhança em função das ordens dos polinômios para os efeitos genéticos aditivo (A), ambiente permanente (P) e dos graus para a regressão fixa (F) e do número de classes de variância residual (H) para característica pldc.

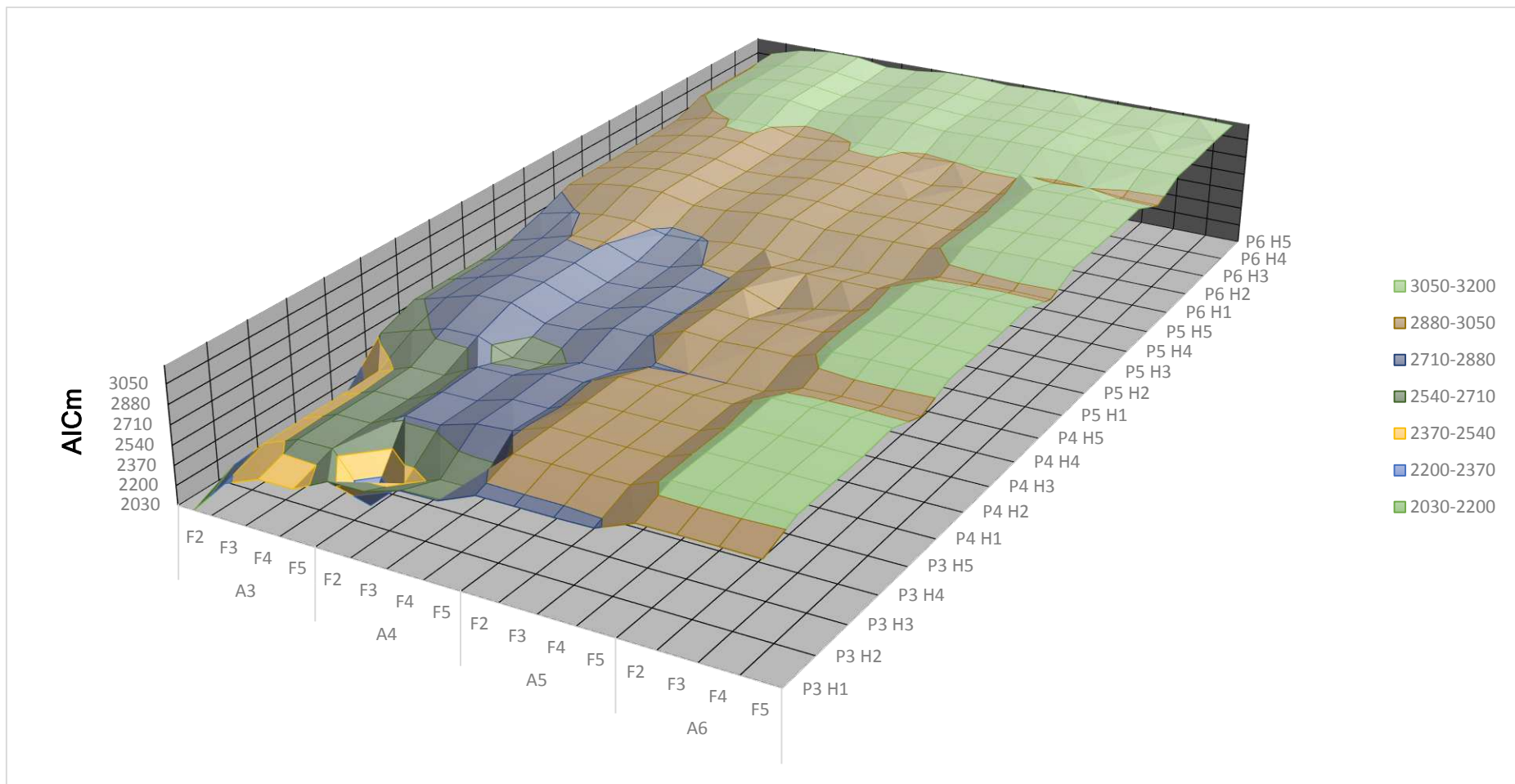


Figura 8: Resultados do critério da informação de Akaike modificado ($AIC_m = -1/2AIC$) em função das ordens dos polinômios para os efeitos genéticos aditivo (A), ambiente permanente (P) e dos graus para a regressão fixa (F) e do número de classes de variância residual (H) para a característica pldc.

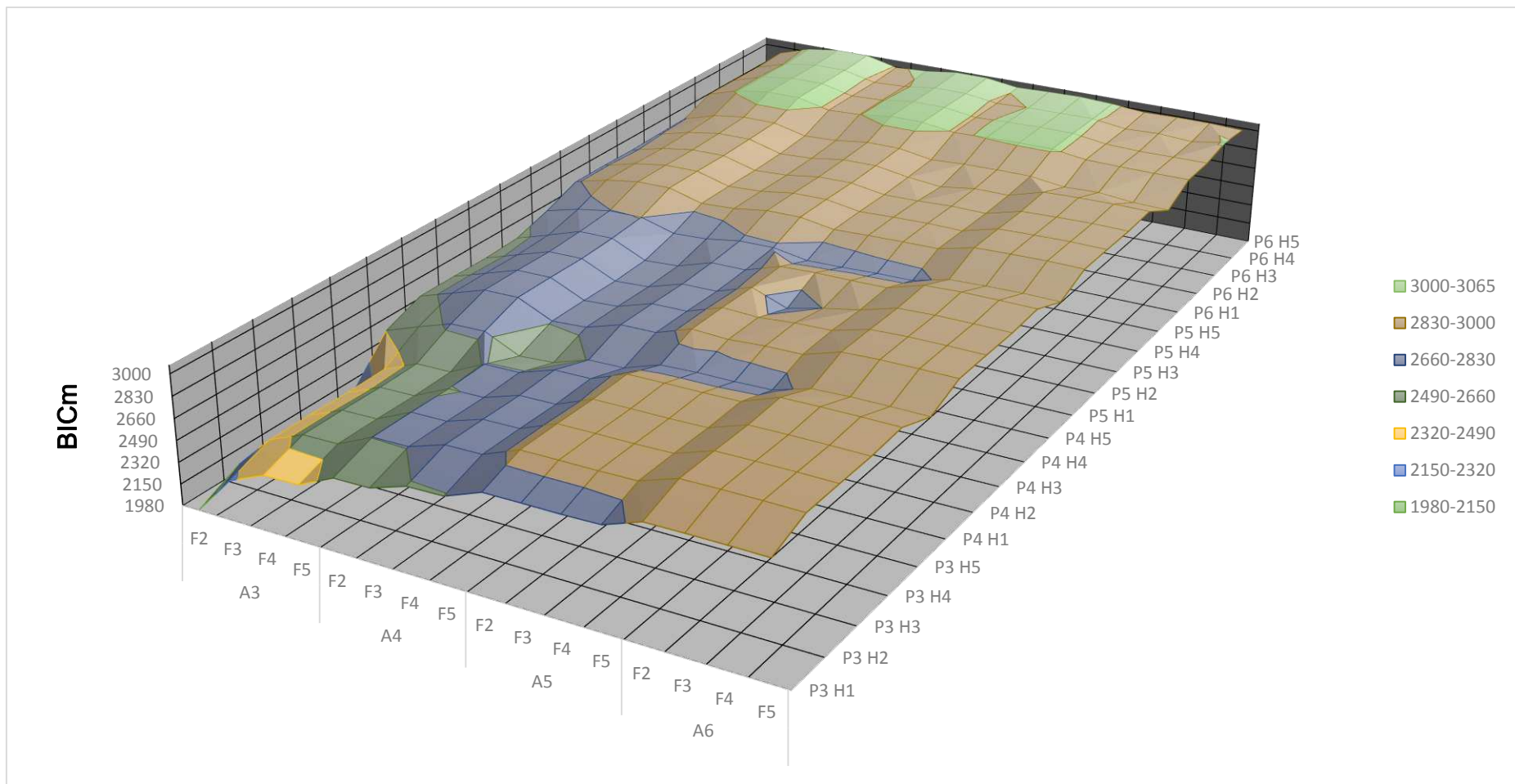


Figura 9: Resultados do critério da informação Bayesiano modificado ($BICm = -1/2BIC$) em função das ordens dos polinômios para os efeitos genéticos aditivo (A), ambiente permanente (P) e dos graus para a regressão fixa (F) e do número de classes de variância residual (H) para característica pldc.

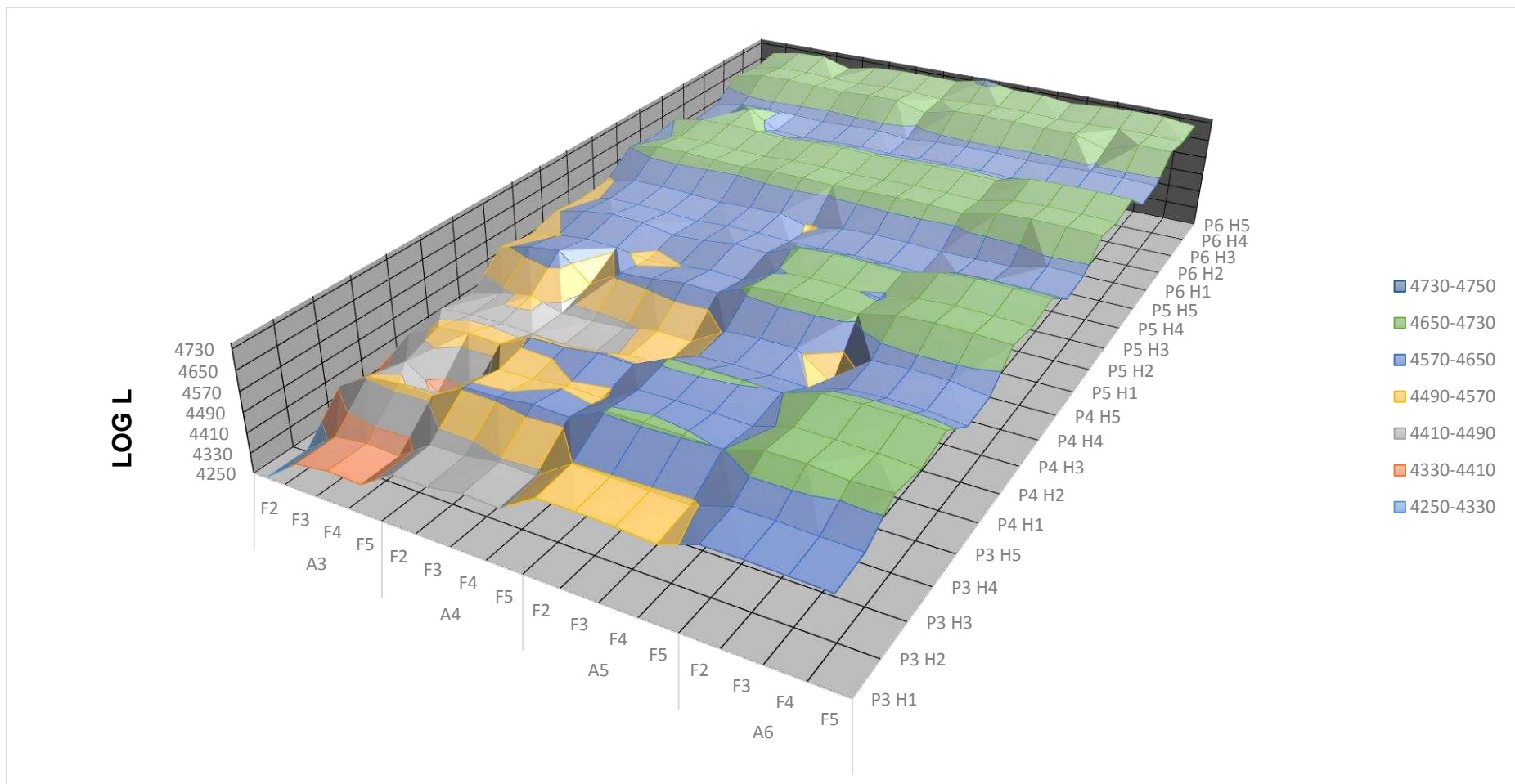


Figura 10: Resultados do logaritmo da função de máxima verossimilhança em função das ordens dos polinômios para os efeitos genéticos aditivo (A), ambiente permanente (P) e dos graus para a regressão fixa (F) e do número de classes de variância residual (H) para característica proteína.

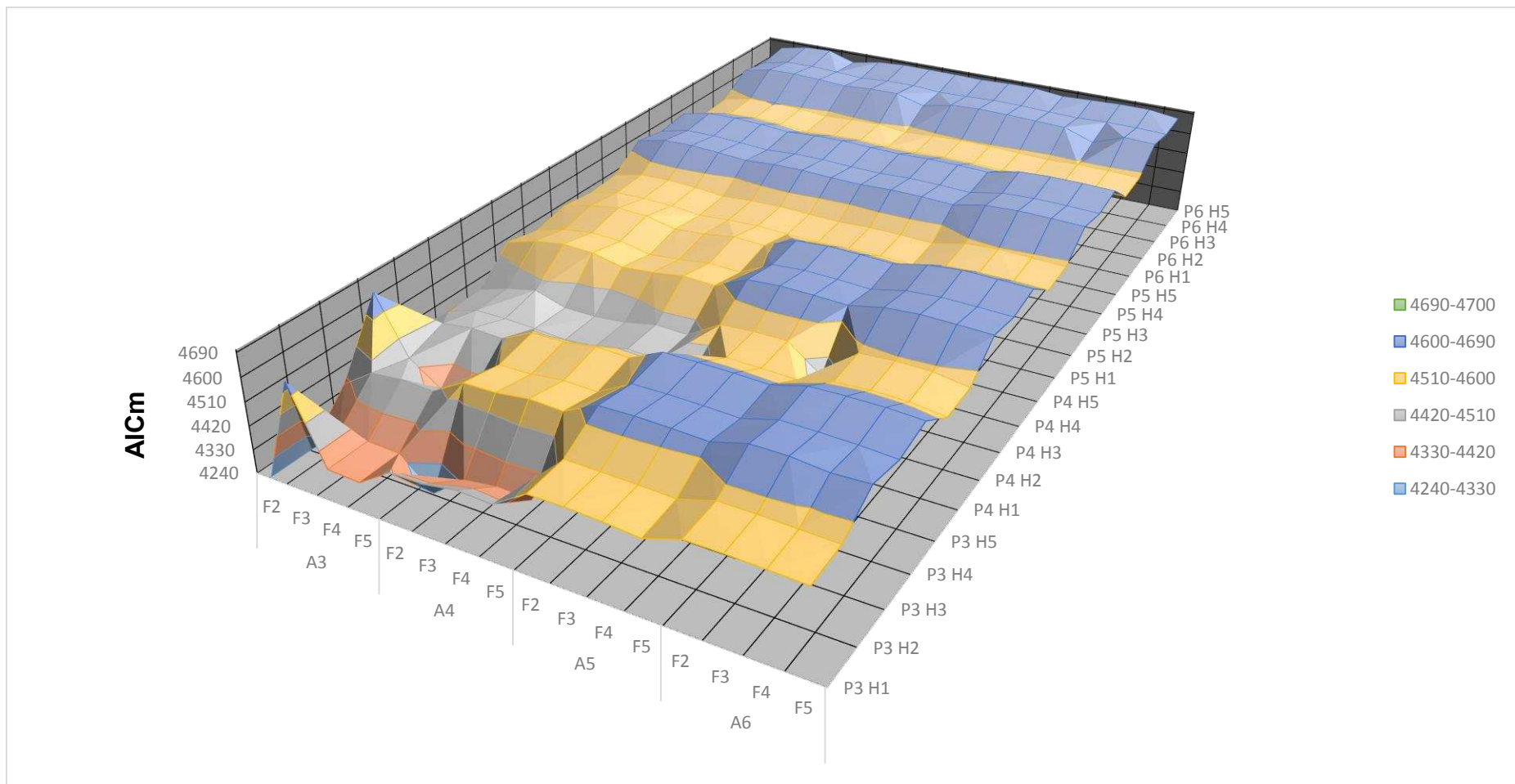


Figura 11: Resultados do critério da informação de Akaike modificado ($AIC_m = -1/2AIC$) em função das ordens dos polinômios para os efeitos genéticos aditivo (A), ambiente permanente (P) e dos graus para a regressão fixa (F) e do número de classes de variância residual (H) para a característica proteína.

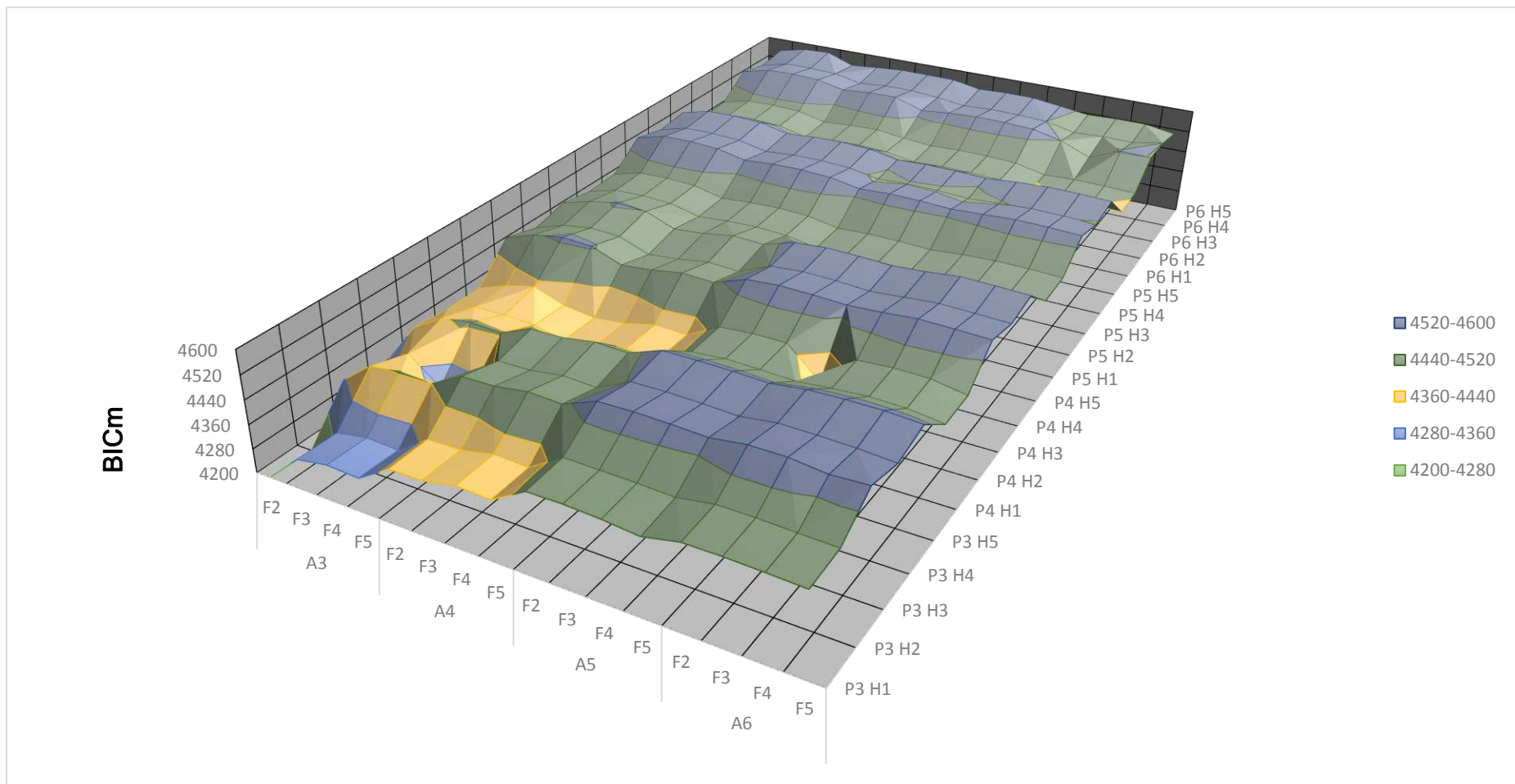


Figura 12: Resultados do critério da informação Bayesiano modificado ($BIC = -1/2BIC$) em função das ordens dos polinômios para os efeitos genéticos aditivo (A), ambiente permanente (P) e dos graus para a regressão fixa (F) e do número de classes de variância residual (H) para característica proteína.

Também é possível observar nas figuras 1, 2, 3, 10, 11 e 12 que as características proteína e gordura apresentaram comportamento instável no momento em que há mudança nas ordens das regressões genética e de ambiente permanente, e para as variâncias residuais. Este comportamento é devido a influência que essas características sofrem no decorrer da lactação do indivíduo pela produção do leite. O que não foi observado para as características lactose e a pldc.

No período de análise foi possível perceber que os modelos com ordens maiores e mais próximas para regressões genética e de ambiente permanente, tiveram maiores dificuldade de convergência, apresentando valores diferentes e que só foi possível alcançar através do algoritmo Px-EM (Parameter Expanded Expectation Maximization) proposto por Meyer (2006), já que são modelos mais parametrizados.

Segundo Foulley e Van Dyk (2000), em modelos de regressão aleatória o algoritmo EM (Expectation Maximization) proposto por Dempster et al. (1977), pode convergir para resultados diferentes de acordo com os valores de partida utilizados, mas algoritmos Px-EM além de rápidos são mais confiáveis.

Dentre os 320 modelos de regressão aleatória avaliados para cada característica, os 7 melhores, de acordo com os maiores valores de AIC e BIC estão apresentados na Tabela 1.

Tabela 1 – Os 7 melhores modelos de cada característica, número de parâmetros (NP), Log L, AICm, BICm e TRV para os polinômios ortogonais de Legendre, considerando heterogeneidade de variância residuais.

Modelo*	NP	Log L	AICm	BICm	TRV
PLDC					
F5A4P6H5	36	3.221,577	3.185,577	3.045,173	5,762 ^{ns}
F5A3P6H5	32	3.217,274	3.185,274	3.060,471	2,844 ^{ns}
F6A4P6H5	36	3.220,914	3.184,464	3.059,662	4,436 ^{ns}
F6A3P6H5	32	3.216,464	3.184,464	3.059,662	4,464 ^{ns}
F5A5P6H5	41	3.221,828	3.180,828	3.020,924	6,264 ^{ns}
F4A5P6H5	41	3.215,764	3.174,764	3.014,859	5,864 ^{ns}
F4A6P6H5	47	3.218,696	3.171,696	2.988,390	Completo

Teor de Gordura

F5A6P3H4	13	-833,722	-846,722	-882,791	14,028 ^{ns}
F5A6P3H3	12	-836,809	-848,809	-882,103	20,202 ^{ns}
F5A6P4H3	16	-782,411	-798,411	-842,804	88,594 ^{**}
F5A6P4H4	17	-832,200	-849,200	-896,367	10,984 ^{ns}
F5A6P5H3	21	-828,750	-849,75	-908,015	4,084 ^{ns}
F5A5P6H3	21	-829,613	-850,613	-908,878	5,81 ^{ns}
F5A6P5H4	22	-826,708	-848,708	-909,748	Completo

Teor de Lactose

F3A4P6H3	34	5.814,617	5.780,617	5.668,248	20,87 ^{ns}
F3A3P6H3	30	5.810,354	5.780,354	5.681,205	29,396 ^{ns}
F3A5P6H3	39	5.818,109	5.779,109	5.650,215	13,886 ^{ns}
F4A3P6H3	30	5.805,513	5.775,513	5.676,367	39,078 ^{**}
F4A6P6H3	45	5.820,276	5.775,276	5.626,556	9,552 ^{**}
F4A4P6H3	34	5.809,877	5.775,877	5.663,511	30,168 ^{**}
F3A6P6H3	45	5.825,052	5.780,052	5.631,328	Completo

Teor de Proteína

F5A3P6H5	32	4.715,945	4.683,945	4.578,192	3,682 ^{ns}
F5A3P6H3	30	4.713,017	4.683,017	4.583,874	9,538 ^{ns}
F4A3P6H3	30	4.709,913	4.679,913	4.580,767	15,746 ^{ns}
F5A4P6H3	34	4.713,671	4.679,671	4.567,308	8,23 ^{ns}
F4A3P6H5	32	4.711,137	4.679,137	4.573,381	13,298 ^{ns}
F4A4P6H3	34	4.712,285	4.678,285	4.565,918	11,002 ^{ns}
F4A5P6H3	39	4.717,786	4.678,786	4.549,895	Completo

** significativo a 1% de probabilidade; *Modelos: FxAxPxHy, em que x corresponde à ordem dos polinômios para curva fixa (F) e para os efeitos aleatórios genético aditivo (A) e de ambiente permanente (P) e y, ao número de classes de variâncias residuais (H). Valores em negrito indicam o melhor modelo com base em Log L, AICm e BICm.

Maiores valores para o Log L sempre serão esperados para os modelos mais parametrizados, o que não foi observado no presente trabalho, por isso outros critérios que permitam a escolha de modelos mais parcimoniosos como os critérios AIC e BIC devem ser utilizados e vêm se tornando padrão em

pesquisas envolvendo regressão aleatória. De acordo com o TRV, não foram encontradas diferenças significativas ($P>0,01$) entre o modelo completo e todos os modelos reduzidos para pldc e teor de proteína, já teor de gordura e lactose somente um e três, respectivamente, foram significativos ($P>0,01$).

Para a escolha do melhor modelo, o presente estudo levou em consideração primeiramente o teste de significância (TRV), posteriormente os critérios de AICm e BICm foram utilizados e por fim o número de parâmetros (NP). Devido ao impasse existente até o momento entre pesquisadores sobre qual critério (AIC e BIC) devem ser utilizados primeiro na determinação de melhor ajuste, o presente estudo levou em consideração os dois critérios e para todas as características dois modelos apresentaram melhores ajuste aos dados.

Pelo TRV todos os modelos para a característica pldc foram semelhantes ao modelo completo e dentre esses modelos, dois apresentaram melhores ajuste de acordo com os critérios AICm e BICm que foram o F5A4P6H5 e o F5A3P6H5, respectivamente. Levando em consideração a eficiência computacional, modelos menos parametrizados são preferíveis e considerando o número de parâmetros como terceiro critério na escolha, o modelo F5A3P6H5 é o recomendado para explicar a variação da característica pldc. Para teor de lactose três modelos foram não significativos ($P<0,01$) em relação ao modelo completo e dentre esses modelos, dois foram os mais adequados de acordo com os critérios AICm e BICm. O modelo F3A4P6H3 foi o que obteve o melhor valor para AICm e para o BICm foi o F3A3P6H3, mas recomenda-se o F3A3P6H3 por apresentar menos NP.

Para a característica proteína, todos os modelos apresentaram o mesmo ajuste que o modelo completo pelo TRV e dentre esses modelos, dois modelos se destacam, o F5A3P6H5 e o F5A3P6H3 por apresentar maiores valores para o AICm e o BICm, respectivamente. O F5A3P6H3 é o recomendado por apresentar menor NP, visto que modelos menos parametrizados são preferíveis. Para teor de gordura somente um modelo foi diferente ao modelo completo pelo TRV e dentre os modelos que apresentaram semelhança de ajuste em relação ao modelo completo, os F5A6P3H4 e o F5A6P3H3 foram os que obtiveram maiores valores para os critérios AICm e o BICm,

respectivamente, mas o modelo F5A6P3H3 foi o escolhido por apresentar menor número de parâmetros.

Componentes de variância e parâmetros genéticos

As variâncias estimadas pelo modelo F5A3P6H5 para característica pldc variaram de 0,6 a 1,0; 0,3 a 0,6; 0,1 a 0,2 e de 0,1 a 0,2 para variância fenotípica, ambiente permanente, residual e genética aditiva, respectivamente (Figura 13). Para teor de gordura as variâncias estimadas para o modelo F5A6P3H3 variou de 0,7 a 2,4; de 0,01 a 1,4; 0,5 a 0,7 e de 0,01 a 0,03 para variância fenotípica, ambiente permanente, residual e genética aditiva, respectivamente (Figura 14). Para teor de lactose as variâncias estimadas no modelo F3A3P6H3 variaram de 0,05 a 0,11; de 0,03 a 0,06; de 0,02 a 0,04 e de 0,01 a 0,04 para variância fenotípica, ambiente permanente, residual e genética aditiva, respectivamente (Figura 15). As variâncias estimadas no modelo F5A3P6H3 para teor de proteína variaram de 0,08 a 0,36; de 0,02 a 0,26; de 0,03 a 0,07 e de 0,04 a 0,06 para variância fenotípica, ambiente permanente, residual e genética aditiva, respectivamente (Figura 16).

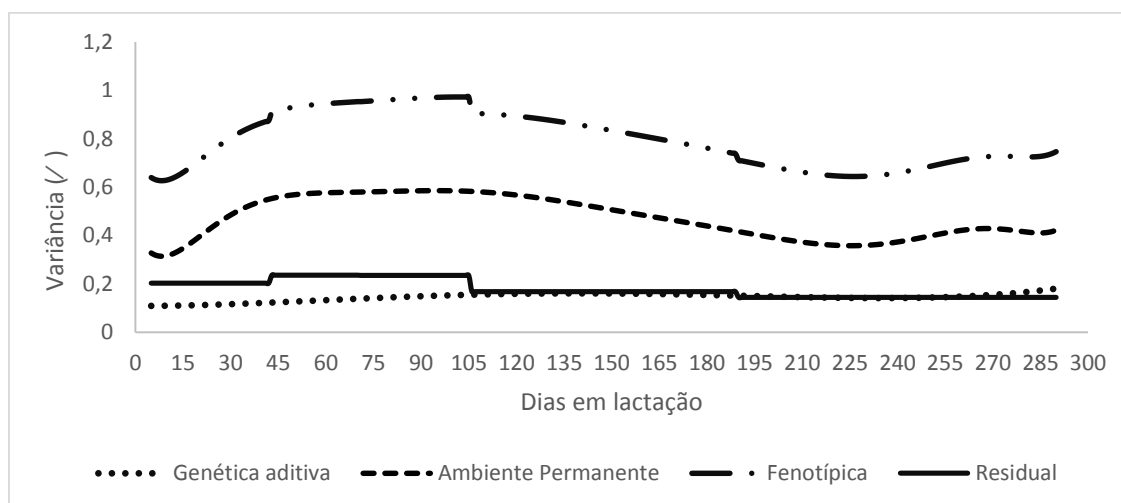


Figura 13: Estimativas das variâncias genética aditiva, ambiente permanente, fenotípica e residual em função dos dias em lactação para pldc para o modelo F5A3P6H5.

Como pode ser observado na Figura 13, as variâncias fenotípicas apresentaram maiores valores entre 45º e 105º dias em lactação, o que é esperado, já que é nesse período que o indivíduo apresenta o pico de produção. A partir dos 105º dias de lactação as variâncias fenotípicas

diminuíram, exceto no final, em que houve ligeiro aumento, provavelmente em razão do menor número de informações.

A variância genética aditiva, aumentou conforme o número de dias em lactação, sendo baixa no início da lactação e alta no final, indicando uma maior variabilidade genética quanto à duração da lactação do que a própria produção em si, uma vez que na primeira lactação costuma-se mensurar a produção de todas as cabras, tendo como consequência essa maior variabilidade. Em relação ao efeito de ambiente permanente, as estimativas foram altas dos 45° aos 135° dias, o que seria esperado por efeito de escala, indicando que nesse período da curva de lactação é muito suscetível a condições do meio.

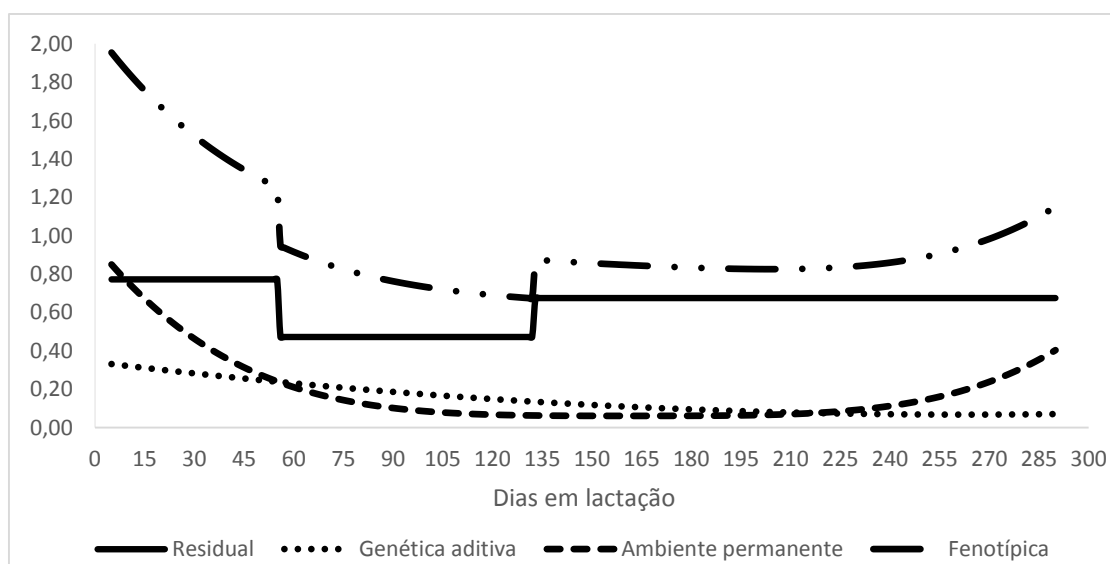


Figura 14: Estimativas das variâncias genética aditiva, ambiente permanente, fenotípica e residual em função dos dias em lactação para teor de gordura para o modelo F5A6P3H3.

Na Figura 14, verifica-se que para teor de gordura, ocorreu diminuição da variabilidade genética no decorrer da lactação, comportamento este inverso apresentado pelas características pldc e proteína.

A variância fenotípica apresentou no início da lactação valores altos e foi diminuindo até os 135° dias de lactação, após esse período observa-se um pequeno aumento. É possível observar que no período do 60° até os 135° foi a fase em que os valores de variância fenotípica apresentaram os menores valores em relação ao todo o período, esse comportamento é esperado já que a uma diminuição no teor de gordura em decorrência da maior produção de leite.

A variância de ambiente permanente apresentou valores baixos no decorrer da lactação. Os altos valores encontrados no início e no final da lactação, provavelmente em razão do menor número de informações.

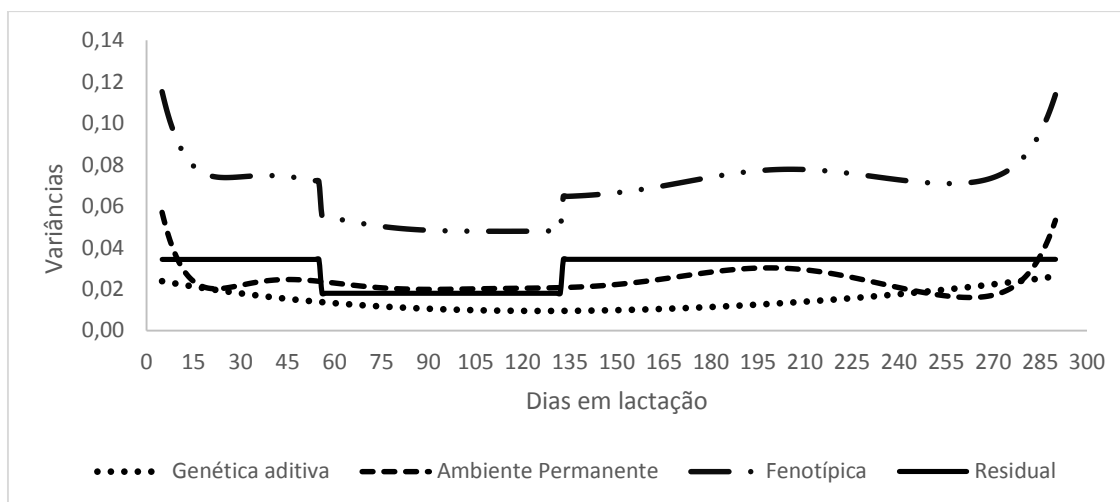


Figura 15: Estimativas das variâncias genética aditiva, ambiente permanente, fenotípica e residual em função dos dias em lactação para teor de lactose para o modelo F3A3P6H3.

Na Figura 15 a variância fenotípica também apresentou o mesmo comportamento apresentado pela característica proteína e gordura, o que mostra que estas características estão diretamente sendo influenciada pela produção de leite. Já a variância genética apresentou valores baixos no decorrer da lactação, sendo que no início e no final foram os períodos em que os valores foram maiores, esses valores podem estar superestimados ou subestimados em razão do menor número de informações no banco de dados.

A variância de ambiente permanente apresentou flutuações nos valores durante os 290 dias de lactação, em que no período do 35º a 60º e do 180º a 210º os valores foram maiores, o que mostra que nestes períodos a influência do meio é maior.

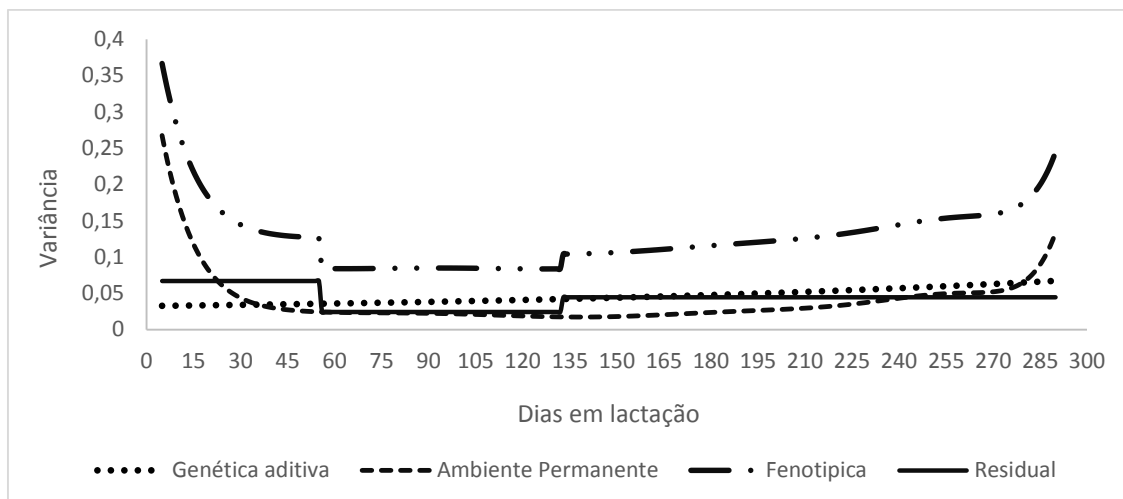


Figura 16: Estimativas das variâncias genética aditiva, ambiente permanente, fenotípica e residual em função dos dias em lactação para teor de proteína para o modelo F5A3P6H3.

As variâncias fenotípicas (Figura 16) apresentaram valores altos no início e queda no período compreendido entre 60° a 135° dias de lactação, comportamento esse esperado já que é nesse período em que o indivíduo apresenta maior produção de leite e levando a uma diminuição no teor de proteína em consequência da diluição pelo leite. Após os 135° dias a variância volta a aumentar os seus valores devido ao período em que as cabras apresentam menores produções, já que estão em final de lactação.

As variâncias de ambiente permanente (Figura 16) durante os dias em lactação apresentaram valores baixos e constantes, o que reforça a ideia de que a característica em estudo sofre pouca influência do meio. Já para a variância genética o comportamento foi semelhante ao apresentado para a característica pldc, em que foram aumentando em decorrência dos dias em lactação o que mostra uma maior variabilidade genética quanto à duração da lactação.

Herdabilidade

Para a característica pldc as estimativas de herdabilidade para o modelo F5A3P6H5 variaram de 0,13 a 0,24 (Figura 17). A herdabilidade apresentou crescimento gradativo no decorrer da lactação, mostrando maiores valores no terço final. Já no início até o 45° a herdabilidade apresentou queda.

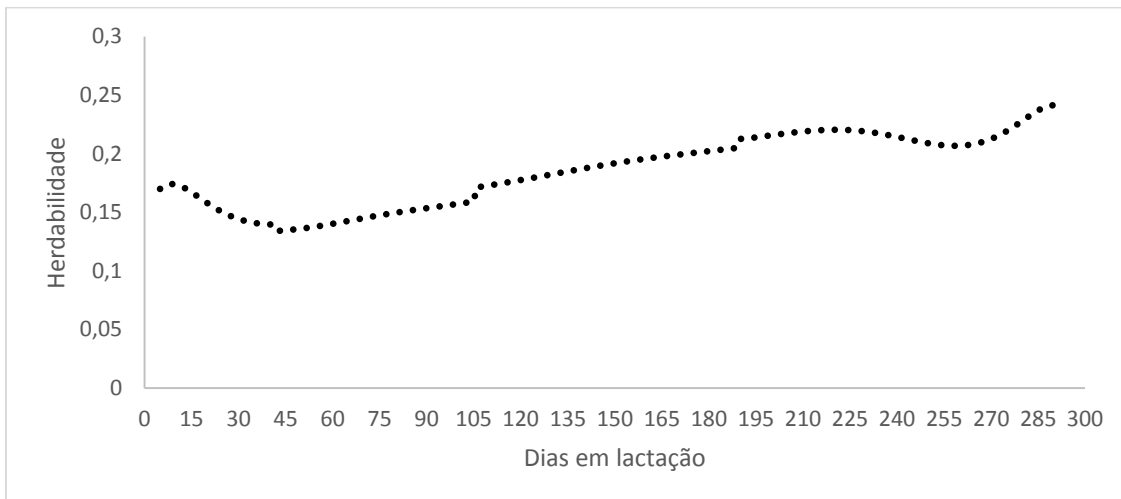


Figura 17: Estimativas de herdabilidade em função dos dias em lactação, para pldc, para o modelo F5A3P6H5.

As estimativas de herdabilidade para a característica gordura variaram de 0,05 a 0,20 para o modelo F5A6P3H3 (Figura 18). Apresentou crescimento do início até os 50º dias em lactação, estimativas de herdabilidade altas no período de 60º a 135º dias em lactação, as quais decresceram até o final da lactação.

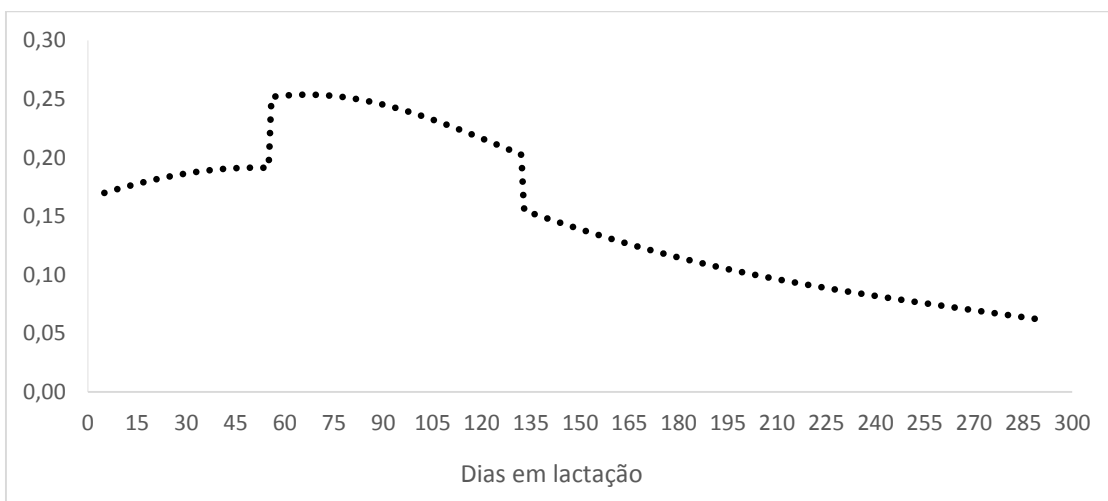


Figura 18: Estimativas de herdabilidade em função dos dias em lactação, para teor de gordura, para o modelo F5A6P3H3.

Para lactose as estimativas de herdabilidade variaram de 0,13 a 0,43 para o modelo F3A3P6H3 (Figura 19). No período de 15º e 50º dias em lactação a herdabilidade apresentou queda, voltando a apresenta valores

maiores no intervalo de 60° até 135° e posteriormente tornando a cair. O final da lactação a herdabilidade apresentou crescimento.

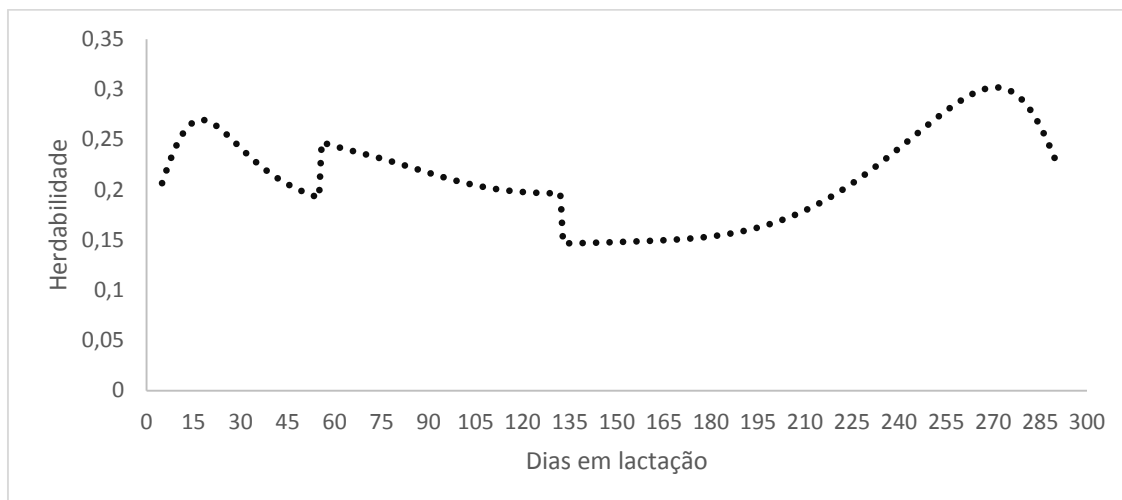


Figura 19: Estimativas de herdabilidade em função dos dias em lactação, para teor de lactose, para o modelo F3A3P6H3.

As estimativas de herdabilidade para proteína variaram de 0,08 a 0,5 para o modelo F5A3P6H3 (Figura 20). Verifica-se que as estimativas apresentaram crescimento até o 135° dias em lactação, sendo que os maiores valores foram encontrados no período de 60° até o 135° e após a esse período a herdabilidade estabilizou-se com leve queda no final.

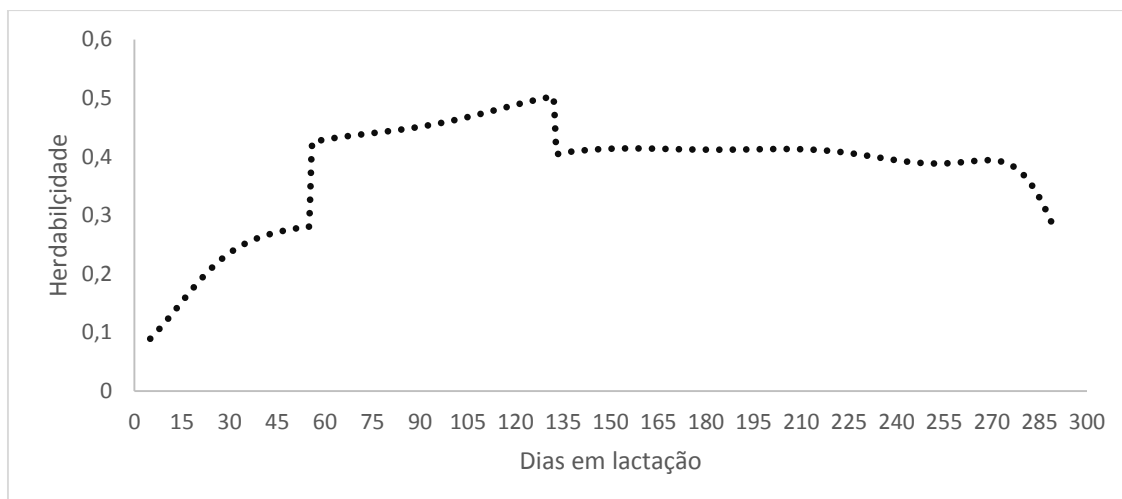


Figura 20: Estimativas de herdabilidade em função dos dias em lactação, para teor de proteína, para o modelo F5A3P6H3.

DISCUSSÃO

Cr terios de avalia o de modelos

Todos os setes modelos para caracter stica pldc foram semelhantes estatisticamente ($P > 0,01$) ao modelo completo pelo TRV e dentre estes modelos o que melhor se ajustou de acordo com o BICm e NP foi o F5A3P6H5. Trabalhando com mesmo rebanho, Sarmiento et al. (2008) sugeriram que um modelo mais parcimonioso, de quarta ordem para o efeito gen tico aditivo e de sexta ordem para ambiente permanente,   suficiente para modelar as mudan as nas vari ncias ao longo da curva de lacta o de cabras. Estes resultados divergem de Silva et al. (2001) trabalhando com cabras da ra a alpina deste mesmo rebanho, em que os autores indicaram que o modelo para curva fixa de quarta ordem, segunda e s tima ordem para o efeito gen tico aditivo e de ambiente permanente, respectivamente, e ao menos cinco classes de vari ncia residuais como melhor modelo, por ser o mais simples e ter convergido facilmente pelo algoritmo AI-REML.

  poss vel notar que assumir cinco classes de vari ncia residual para caracter stica pldc em caprinos de ra a alpina evidencia que a caracter stica   influenciada por muitos fatores e que podem diferir ao longo da lacta o. Estes resultados est o de acordo com os apresentados por Silva et al. (2011) e Brito (2012).

Para a caracter stica gordura um modelo foi significativo ($P > 0,01$), portanto o modelo completo fornece melhor ajuste. Para os modelos que foram semelhantes ao modelo completo, o que melhor se ajustou de acordo com crit rio BICm e NP foi o F5A6P3H3. Zampar (2012) trabalhando com modelos de regress o aleat ria em bovinos leiteiros da ra a holandesa, recomenda que polin mios ortogonais de Legendre de quinta e sexta ordem para os efeitos aleat rios, foram os que melhor se adequaram aos teores de gordura. Biassus et al. (2010), trabalhando com MRA para vacas da ra a holandesa tamb m conseguiram melhores ajustes com modelos utilizando sexta ordem.

Tr s modelos para teor de lactose foram semelhantes ao modelo completo pelo TRV e o que melhor se ajustou foi o F3A3P6H3 de acordo com

os critérios BICm e NP. Como estudos com teor de lactose do leite é escasso, fazer comparações com estudos na literatura não foi possível.

Se observamos os modelos referente a curva para o efeito genético aditivo e de ambiente permanente para característica pldc e lactose, as ordens de ajuste foram semelhantes. Esta semelhança, acredita-se que é devido a lactose está diretamente ligada a síntese do leite. Portanto, mais estudos são necessários para definir um modelo que estude o comportamento do teor de lactose ao longo da curva de lactação.

Todos os modelos para a característica proteína foram semelhantes ao modelo completo estatisticamente ($P > 0,01$) pelo TRV e o que melhor se ajustou foi o F5A3P6H3 de acordo com o BICm e NP. Trabalhando com vacas leiteiras da raça holandesas, Zampar (2012) sugeriu que modelo de quarta ordem foi o que melhor se ajustou aos dados.

Estudos com teores de gordura e proteína no leite em caprinos são muito escassos na literatura, já para bovinos leiteiros já é possível encontrar um pouco mais. Porém, vale ressaltar que modelos de regressão aleatória utilizando-se os polinômios ortogonais de Legendre de quarta, quinta e sexta ordem para efeitos aleatórios têm sido comumente selecionados pelos testes de comparações de modelos em estudos com produção de leite em bovinos (Biassus et al., 2010; Costa et al. 2005; Cobuci et al., 2006; Liu et al., 2006).

Componentes de variância e parâmetros genéticos

As estimativas de variância genética aditiva para teor de proteína para o melhor modelo foram baixas (entre 0,04 a 0,06), o que evidencia que a característica apresenta pouca influência genética, mesmo tendo pouca influência é possível observar a existência de variabilidade genética ao longo da lactação. Já os resultados indicam maior variância de ambiente permanente no início e no final da lactação, evidenciando que o estado do animal pós parto influencia na característica e que no decorrer da lactação a característica sofre pouco efeito de ambiente. Se observamos, os menores valores de variância fenotípica estão compreendidos no período de 60º até 135º dias de lactação,

isso é devido a alta produção de leite que o animal apresenta, levando uma diminuição no teor de proteína em consequência da diluição da produção de leite.

Para teor de gordura as variâncias genéticas foram baixas ao longo da lactação. Mesmo tendo pouca variabilidade, é possível perceber que o teor de gordura apresenta maiores influencia genética do que teor de proteína. Já as variâncias de ambiente permanente, fenotípica e residual apresentou o mesmo comportamento que a característica teor de proteína. A gordura é o constituinte mais variável, o que reforça o resultado dos valores para as variâncias (genética, ambiente e fenotípica) serem maiores do que a característica proteína. Comportamento semelhantes a este estudo foram encontrados por Zampar (2012) e Biassus et al. (2009).

Os resultados para pldc indicam maior variância de ambiente permanente no período do 30º a 135º, que seria esperado por efeito de escala. Por outro lado, a variância genética foi aumentando de forma gradativa ao longo da lactação. Esses resultados corroboram com Silva et al. (2011) que utilizou MRA em cabras da raça alpina. Este aumento gradativo indica uma maior variabilidade genética quanto à duração da lactação do que a própria lactação em si, uma vez que na primeira lactação costuma-se mensurar a produção de todas as cabras.

A lactose tem um importante papel na síntese do leite. Sua baixíssima amplitude de variação se deve ao fato de a lactose estar relacionada à regulação da pressão osmótica na glândula mamária, de forma que maior produção de lactose determina maior produção de leite, com o mesmo teor de lactose (González et al. 2001). Diante do papel que a lactose tem na produção do leite, o que elucida a questão dessa característica apresentar valores baixos de variância para componente.

As estimativas de herdabilidade para teor de proteína do modelo variaram de 0,08 a 0,5 e para teor de gordura variaram de 0,05 a 0,20. Esses resultados possuem de baixa a alta herdabilidade, sugerindo que o uso dos mesmos em seleção genética pode ocasionar ganhos genéticos desejáveis na população, quando utilizados nos estágios em que as estimativas são maiores,

o que significa que nesses estágios há maior influência de fatores genéticos. Resultados semelhantes também foram encontrados por Biassus et al. (2009) trabalhando com vacas holandesas, as quais variaram entre 0,09 e 0,33. Já Paula et al. (2008) trabalhando com abordagem Bayesiana encontrou herdabilidade de 0,58 para proteína estimadas em rebanho holandês do Estado do Paraná, já para gordura o resultado foi diferente, com valor de 0,60.

A herdabilidade estimada para teor de gordura pelo modelo estudado apresentou a mesma tendência descrita para as estimativas da variância genética (Figura 18), ou seja, estimativa de herdabilidade apresentou queda gradativa a parti do 60º dia em lactação.

Para teor de lactose as estimativas de herdabilidade variaram de 0,13 a 0,43, de baixa a média magnitude. Este resultado mostra que pode ser utilizada com objetivos de ganhos genéticos desejáveis na população, quando utilizado em estágio em que as estimativas são maiores.

A herdabilidade para produção de leite apresentou comportamento semelhante ao da variância genética aditiva, sendo maior no terço final da lactação, contrastando com os resultados encontrados por Menezes et al. (2010), que encontraram maiores valores de variância genética e herdabilidade no início da curva de lactação e maiores valores de ambiente permanente no final da lactação, utilizando o modelo F3A4EP6. Os resultados encontrados nesse estudo estão de acordo com os apresentados por Silva (2011) em caprinos da raça Alpina, que também encontraram maiores valores de variância genética e herdabilidade no terço final da lactação.

CONCLUSÕES

Modelos de regressão aleatória foram adequados para descrever variações na produção do leite, no teor de proteína, no teor de gordura e no teor de lactose de cabras da raça Alpinas. Dentre os modelos estudados, o mais indicado para avaliação genética para pldc foi o que considerou ordem 5, 3 e 6 para a curva fixa e para os efeitos aleatórios, genético aditivo e de ambiente permanente, respectivamente, e ao menos 5 classes de variância

residual. Para teor de proteína o mais indicado foi curva fixa de ordem 5, a curva de efeito genético de ordem 3, a curva de efeitos de ambiente permanente de ordem 6 e 3 classes de variância residual. Para gordura o modelo de ordem 5, 6 e 3 para curva fixa e para os efeitos aleatórios, genético aditivo e ambiente permanente, respectivamente, e 3 classes de variância residual. Para lactose o mais indicado foi o que considerou a ordem 3, 3 e 6 para efeito fixo e para os efeitos aleatórios, genético aditivo e de ambiente permanente, respectivamente, e 3 classes de variância residual.

A identificação de animais com maior teor de proteína é mais eficiente a partir do 60º dia em lactação, para teor de gordura até 120º dia em lactação, para teor de lactose no início e no final da lactação e para produção de leite a partir do 120º dia em lactação, período em que a expressão da variância aditiva e as herdabilidades indicam que a seleção resultaria em ganhos mais efetivos.

REFERÊNCIA BIBLIOGRAFICA

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

ARAÚJO, C.V.; TORRES, R.A.; COSTA, C.N.; 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.

BIASSUS, I.O. **Modelo de regressão aleatória na estimação de parâmetros genéticos para produção e persistência nas características produtivas de vacas da raça holandesa**. Porto Alegre, UFRGS, 2009. 104p. Dissertação (Mestrado) – Universidade Federal do Rio Grande do Sul, 2009.

BIASSUS, I.O.; COBUCI, J.A.; COSTA, C.N.; RORATO, P.R.N.; BRACCINI NETO, J.; CARDOSO, L.L. Persistence in milk, fat and protein production of primiparous Holstein cows by random regression models. **Revista Brasileira de Zootecnia**, Viçosa, v.39, n. 12, p. 2617-2624, 2010.

BRITO, L.F. Modelagem da produção de leite de cabras das raças **Alpina e Saanen utilizando Regressão Aleatória**. UFV, 2012. 82f. Dissertação

(Mestrado), Programa de Pós-Graduação em Genética e Melhoramento (PPGGM), Universidade Federal de Viçosa (UFV). 2012.

COBUCI, J.A.; COSTA, C.N.; TEIXEIRA, N.M. et al. Utilização dos polinômios de Legendre e da função de Wilkink em avaliações genéticas para persistência na lactação de animais da raça Holandesa. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.58, n.4, p.614-623, 2006.

COSTA, C.N.; MELO, C.M.R. de; MACHADO, C.H.C. et al. Parâmetros genéticos para a produção de leite de controles individuais de vacas da raça Gir estimados com modelos de repetibilidade e regressão aleatória. **Revista Brasileira de Zootecnia**, v.34, n.5, p.1519-1530, 2005.

DEMPSTER, A.P.; LAIRD, N.M. and RUBIN, D.B. 1977. Maximum likelihood from incomplete data via the EM algorithm. **J. Roy. Stat. Soc. B** 39:1-38.

FOULLEY, J.L.; VAN DYK, D.A. The PX-EM algorithm for fast stable fitting of Henderson's mixed model. **Genet. Sel. Evol**, v.32, p.143-163. 2000.

GONZÁLEZ, F. H. D. Perfil metabólico en bovinos: alcance y utilidad. **Revista MVZ**, v. 3, n.3, p. 45-52, 2001.

HENDERSON Jr., C.R. Analysis of covariance in the mixed model: higher-level, nonhomogeneous, and random regressions. **Biometrics**, Washington, v.38, p.623-640, 1982.

KIRKPATRICK, M.; LOFSVOLD, D.; BULMER, M. Analysis of the inheritance, selection and evolution of growth trajectories. **Genetics**, v.24, n.3, p.979-993, 1990.

LIU, Y.X.; ZHANG, J.; SCHAEFFER L.R. et al. Short communication: Optimal random regression models for milk production in dairy cattle. **Journal of Dairy Science**, Champaign, v.89, p.2233-2235, 2006.

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.; CAETANO, G.C.; BRITO, L.F.; SILVA, F.G.; MELO, A.L.P.; TORRES, R.A.; RODRIGUES, M.T. **Estudo da produção de leite em cabras da raça Saanen utilizando modelos de regressão aleatória**. VIII Simpósio Brasileiro de Melhoramento Animal Maringá, PR – 01 e 02 de julho de 2010.

MEYER, K. **PX x AI: Algorithmic for better convergence in restricted maximum likelihood estimation**. Proceedings of the Eighth World

Congresso in Genetic Applied to Livestock Production, communication no. 24-15. 2006.

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.

PAULA, M.C.; MARTINS, E.N.; SILVA, L.O.C.; OLIVEIRA, C.A.L.; VALOTTO, A.A.; GASPARINO, E. Estimativas de parâmetros genéticos para produção e composição do leite de vacas da raça Holandesa no estado do Paraná. **Revista Brasileira de Zootecnia**, Viçosa, v. 37, p. 824-828, 2008.

RAO, C.R. **Linear statistical inference and its applications**. 2.ed. New York: John Willy & Sons. p.417-420. 1973.

SARMENTO, J.L.R.; ALBUQUERQUE, L.G; TORRES, R.A. et al. Comparação de modelos de regressão aleatória para estimação de parâmetros genéticos em caprinos leiteiros. **Revista Brasileira de Zootecnia**, v.37, p.1788-1796, 2008.

SCHWARZ, G. Estimating the Dimension of a Model. **Annals of Statistics**, 1978, 6, p.461-464.

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.

ZAMPAR, A. **Modelos de regressão aleatória para característica de qualidade de leite bovino**. Piracicaba, ESALQ, 2012. 52p. Tese (doutorado) – Universidade de São Paulo, 2012.

CAPÍTULO II

Avaliação genética da produção e dos constituintes do leite de caprinos da raça Alpina utilizando modelos de regressão aleatória multicaracterísticos

RESUMO: Este estudo teve como objetivo verificar se os modelos de regressão aleatória multicaracterístico (MRAM) podem ser aplicados adequadamente na avaliação da produção e dos constituintes do leite no dia do controle de cabras da raça alpina. Para isso utilizou-se 34.635 informações das características produção de leite, teor de proteína, gordura e lactose de 644 cabras da raça alpina de primeira lactação. Para a escolha dos modelos que melhores se ajustaram aos dados, os seguintes critérios foram utilizados: logaritmo do máximo da função de verossimilhança (Log L), critério da informação de Akaike (AIC), critério da informação Bayesiano (BIC), teste da razão de verossimilhança (TRV) e número de parâmetros (NP). O modelo que considerou ordem três para a curva fixa, ordem três para o efeito genético aditivo, ordem cinco para o efeito de ambiente permanente e três classes de variância residual foi o mais indicado para avaliação genética das quatro características estudadas, por apresentar o melhor valor de AICm ($-1/2 \times AIC$) e menor número de parâmetros. A variância genética aditiva para produção de leite no dia do controle (pldc) e teor de proteína aumentava conforme os dias de lactação iam aumentando, já para teor de gordura e lactose os maiores valores só foram encontrados nos extremos da curva e no meio as estimativas foram baixas. As estimativas de herdabilidade variaram de 0,18 a 0,23; de 0,09 a 0,45; de 0,06 a 0,25 e de 0,13 a 0,26 para as características pldc, proteína, gordura e lactose, respectivamente. Os MRAM apresentam maiores dificuldade de convergência em relação aos modelos de regressão aleatória unicaracterístico, visto que são modelos muito parametrizados.

Palavras-chave: genético aditivo, gordura, lactose, proteína

Genetic evaluation of production and constituents of milk from Alpine goats
using multiple-trait random regression models.

ABSTRACT: This study aimed to verify if multiple-trait random regression models (RRMM) can be properly applied in evaluation the production and constituents traits in dairy goats control of the Alpine race. For this we used information 34,635 milk production, protein, fat and lactose from 644 goats alpine breed from first lactation. For the selection of models that best adjusted to the data, the following criteria were used: maximum of likelihood function (log L), Akaike information criterion (AIC), Bayesian Information Criterion (BIC), ratio test likelihood (TRV) and number of parameters (NP). The model that showed the best fit was from order three for the fixed curve, order three for the additive genetic effect, order five for the permanent environmental effect and four residual variance classes for the genetic evaluation for the four traits by presenting the best AICM ($-1/2 \times AIC$) and fewer parameters. The additive variance for yield milk test-day (TD) and protein content increased as the days of lactation were increasing, as for fat and lactose the highest values were found only in extreme curve and in the middle estimates were low. Heritability estimates ranged from 0.18 to 0.23; 0.09 to 0.45; 0.06 to 0.25 and 0.13 to 0.26 for milk yield characteristics, protein, fat and lactose, respectively. The RRRM present greater convergence difficulties in relation to simple trait random regression models, as they are very parameterized models.

Keywords: additive genetic, protein, fat, lactose

INTRODUÇÃO

Registros de produção de leite e seus componentes medidos ao longo da lactação fornecem um exemplo típico de medições repetidas ou dados longitudinais (medições em sequência obtida a partir de um mesmo animal ao longo do tempo). A principal característica dos dados longitudinal é a avaliação das propriedades genéticas de uma ou mais variáveis através de uma dimensão específica, como o tempo ou o espaço (Wiggans e Goddard, 1997).

Vários estudos mostraram que os modelos de regressão aleatória (MRA) são uma abordagem atraente para modelar medidas repetidas e estimar parâmetros genéticos para estas características (Schaeffer e Jamrozik, 2008; Sesana et al., 2010), pois permite aproveitar melhor informações, por ser menos sensível a dados perdidos, modelar as variações de produção dentro da lactação e trabalhar com número maior de dados. Em contraste com os modelos multicaracterísticos que fornecem previsões pontuais, o modelo de regressão aleatória permite a predição de valores genéticos para a curva de lactação como um todo, para qualquer ponto ou intervalo dentro do período de lactação, e para funções da curva de lactação (Jamrozik et al., 1997). Em gado leiteiro, as regressões são providas como uma função do período de lactação, uma função de base que permite a modelação da trajetória tanto para a média da população (regressão fixo) e de cada animal (regressão aleatória). Outra vantagem de MRA é que a estrutura de covariância entre as medidas pode ser caracterizada por uma função, um fato que permite uma maior flexibilidade na utilização de dados longitudinais (Van der werf e Schaeffer, 1997).

Embora, vários estudos relatem estimativas de parâmetros genéticos para registros de produção de leite no dia do controle e seus componentes, obtidos por MRA em bovinos de leite (Schaeffer e Jamrozik, 2008) e búfalos (Breda et al., 2010; Sesana et al., 2010; Aspilcueta-Borquis et al., 2010), poucos estudos se aplicam a modelos de regressão aleatória multicaracterístico (MRAM) para as característica produção de leite no dia do controle, de gordura e de proteína (Muir et al, 2007; Hammani et al., 2008).

Ao se avaliar mais de uma característica de forma individual utilizando MRA traz algumas desvantagens, uma delas é a não possibilidade de

aproveitar simultaneamente as informações de mais de uma característica e a estrutura de covariância entre as características. A implementação de MRAM é atraente quando se trabalha com mais de uma característica, pois permite aproveitar melhor as informações ao estudar duas ou mais características simultaneamente, o que aumenta o volume de informação e permite aproveitar a estrutura de covariância entre as características, obtendo previsões muito mais acuradas para os valores genéticos em relação aos MRAU.

Em comparação com o gado leiteiro, a produção total de leite de cabra é menor, a duração da lactação é mais curta, e o pico de lactação é menos acentuado. Tendo em vista a possibilidade de usar MRA e a importância dos componentes do leite, são necessários estudos que investiguem a associação entre o leite no dia do controle e a produção de gordura, proteína e lactose ao longo da lactação.

Objetivou-se com este estudo verificar se os MRAM podem ser aplicados adequadamente na avaliação da produção e dos constituintes do leite no dia do controle de cabras da raça Alpina com a finalidade de definir o modelo mais adequado e estimar componentes de (co)variância e os parâmetros genéticos.

MATERIAL E MÉTODOS

Os registros de produção e constituintes do leite de caprinos da raça Alpina utilizados neste estudo foram provenientes do setor de caprinocultura do Departamento de Zootecnia (DZO) da Universidade Federal de Viçosa (UFV) em Viçosa, Minas Gerais. A produção de leite no dia do controle (PLDC) que representa a soma do leite coletado no período de 24 horas foi obtida por meio de controles leiteiros realizados semanalmente em ordenha mecânica, sendo duas vezes ao dia. A primeira coleta teve início às seis horas da manhã e a segunda coleta teve início às quatorze horas. Após o término do controle leiteiro de cada semana, os dados eram conferidos e digitados no banco de dados.

Os constituintes do leite utilizado foram obtidos a partir de amostras coletadas durante o período de 1999 à 2014, sendo realizada uma vez ao mês em recipiente de aproximadamente 75 ml, em que no período da manhã eram coletados a fração de 2/3 e a tarde eram coletados 1/3 do volume total, feito isso, os recipientes recebiam identificações das cabras e armazenados com conservantes e posteriormente enviados ao laboratório de qualidade do leite do DZO, que retornava os dados de gordura, proteína e lactose do leite, além da contagem de células somáticas. Estes dados eram conferidos e digitados no banco de dados.

A organização e preparação do arquivo de dados foram realizadas utilizando-se o programa SAS 9.1. Após análise crítica dos dados restaram 34.635 informações de 644 cabras da raça Alpina de primeira lactação. As lactações foram truncadas aos 290 dias e animais que tiveram o primeiro parto entre os 7 e 36 meses de idade foram mantidos na análise.

Para corrigir erros de pedigree e recodificar os dados alfanuméricos foi utilizado o programa RENPED (Silva, 2011). Já o programa WOMBAT (Meyer, 2007), foi utilizado em todas as análises genéticas, fornecendo as estimativas dos parâmetros para cada animal avaliado e bem como os valores necessários para o cálculo dos critérios de escolha dos modelos.

As análises foram realizadas para as características pldc, proteína, gordura e lactose. No arquivo, apenas foram utilizados animais com composição genética (Alpina = 100% = 1; $90\% \leq \text{Alpina} < 100\%$ = 2; $80\% \leq \text{Alpina} < 90\%$ = 3; $70\% \leq \text{Alpina} < 80\%$ = 4; $60\% \leq \text{Alpina} < 70\%$ = 5; $50\% \leq \text{Alpina} < 60\%$ = 6). O ano-estação do controle (de 2000 a 2014 e estação 1 de março a setembro e 2 para os demais meses), tipo de parto (até três filhotes) e idade da cabra ao parto como covariável, considerando-se os efeitos linear e quadrático.

O modelo animal multicaracterístico de regressão aleatória utilizando nas avaliações simultâneas da pldc, proteína, lactose e gordura é uma extensão dos modelos de regressão aleatória unicaracterístico (MRAU) apresentado no primeiro capítulo desse mesmo estudo.

O modelo representado sob forma matricial é:

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

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z & X'R^{-1}W \\ Z'R^{-1}X & Z'R^{-1}Z + A^{-1} \otimes G^{-1} & Z'R^{-1}W \\ W'R^{-1}X & W'R^{-1}Z & W'R^{-1}W + I \otimes P^{-1} \end{bmatrix} \cdot \begin{bmatrix} b \\ a \\ p \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \\ W'R^{-1}y \end{bmatrix}$$

em que:

y é o vetor de observações (pldc, proteína, gordura e lactose); X , Z e W são, respectivamente matrizes de incidência de efeito fixo (grupo contemporâneo e a covariável), efeito aleatório genético aditivo e de ambiente permanente; b , a e p vetores de efeitos fixos, genéticos aditivos e ambiente permanente; e vetor de erros aleatório.

Este modelo compreende os seguintes pressupostos:

$$Var = \begin{bmatrix} a \\ p \\ e \end{bmatrix} = \begin{bmatrix} A \otimes k_d & 0 & 0 \\ 0 & I_d \otimes k_p & 0 \\ 0 & 0 & R \end{bmatrix}$$

em que:

a é o vetor de coeficiente de regressão aleatória genética aditiva;

p é o vetor de coeficiente de regressão aleatório de ambiente permanente;

e é o vetor aleatório residual;

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

k_d e k_p é a matrizes de (co)variâncias entre os coeficientes de regressão aleatórios genético aditivo e de ambiente permanente, respectivamente;

I_d é a matriz identidade de dimensão d (número de animais com registros);

R é a matriz diagonal de variância residual;

\otimes é o operador produto de Kronecker.

Para a escolha dos modelos que melhores se ajustaram aos dados, os seguintes critérios foram utilizados na seguinte ordem: teste da razão de verossimilhança (TRV), critério da informação de Akaike (AIC) (Akaike, 1973) e critério da informação Bayesiano (BIC) (Schwarz, 1978), número de parâmetros (NP) e logaritmo do máximo da função de verossimilhança (Log L).

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 melhores ajuste do modelo. Os critérios de informação podem ser representados como: $AIC = -2 \text{Log } L + 2p$ e $BIC = -2 \text{Log } L + p \log (n - r)$,

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 Log L é o logaritmo do máximo da função de verossimilhança.

Vale ressaltar que os AICs e BICs que serão apresentados neste trabalho estão modificados (AICm e BICm) através da multiplicação pelo termo (-1/2), pois isso faz com que estes critérios fiquem em escala semelhante ao Log L e que o maior valor corresponda ao melhor modelo, o que facilita a comparação dos modelos e comportamento dos critérios frente a alterações nos modelos.

Para comparação entre modelos completos e reduzido também foi utilizado o Teste da Razão de Verossimilhança (TRV) (RAO, 1973), ao nível de 1% de probabilidade. $TRV = -2 [\ln L1 - \ln L2]$

O TRV compara as diferenças entre o máximo das funções de verossimilhança linearizadas de cada modelo, com o valor obtido de uma distribuição qui-quadrado com o número de grau de liberdade igual à diferença no número de parâmetros entre os modelos.

O número de parâmetros também foi utilizado na escolha do modelo, haja vista que modelos menos parametrizados são preferíveis devido ao menor tempo de análise e maior facilidade de convergência.

Quanto a utilização das estruturas de variância residuais seguiu a mesma utilizada para o primeiro capítulo.

RESULTADO

Dentre os 320 modelos de regressão aleatória multicaracterístico avaliados, os 7 melhores, de acordo com os maiores valores de AICm e BICm estão apresentados na Tabela 1.

Tabela 1 – Os sete melhores modelos, número de parâmetros (NP), Log L, AICm, BICm e TRV para os polinômios ortogonais de Legendre, considerando heterogeneidade de variância residuais para regressão aleatória multicaracterístico.

Modelo*	NP	Log L	AICm	BICm	TRV
F2A5P4H1	356	12.830,310	12.474,314	11.970,52	164,782**
F3A3P5H4	328	13.196,445	12.868,445	11.482,95	567,480**
F4A4P3H5	264	12.805,620	12.541,617	11.426,48	214,176**
F4A4P3H4	254	12.736,920	12.482,917	11.426,48	351,576**
F5A3P4H4	254	12.740,731	12.486,731	11.413,84	343,948**
F3A3P5H3	318	12.886,798	12.568,798	11.225,54	51,814 ^{ns}
F2A5P5H1	430	12.912,705	12.482,705	10.666,33	Completo

**Significativo a 1% de probabilidade; *Modelos: FxAxPxHy, em que x corresponde à ordem dos polinômios para curva fixa (F) e para os efeitos aleatórios genético aditivo (A) e de ambiente permanente (P) e y, ao número de classes de variâncias residuais (H). Valores em negrito indicam o melhor modelo com base em Log L, AICm e BICm.

Dentre os modelos de regressão aleatória multicaracterístico analisados (Tabela 1) somente o modelo F3A3P5H3 foi não significativo ($P > 0,01$) em comparação ao modelo completo pelo TRV. Portanto o modelo F3A3P5H3 é o recomendado para explicar as variações das características simultaneamente por apresentar o mesmo ajuste que o modelo completo e por ter menor número de parâmetros. Rodar modelos que envolve muitas características simultaneamente exige alta eficiência computacional, muito tempo de análise e dificuldade de convergência, já que são modelos muito parametrizados.

Componentes de variância e parâmetros genéticos

As variâncias estimadas para as quatro características são provenientes do mesmo modelo (F3A3P5H3). As variâncias para pldc utilizando o modelo de regressão aleatória multicaracterístico variaram de 0,18 a 0,29; de 0,47 a 0,79; de 0,15 a 0,22 e de 0,84 a 1,24 para genética aditiva, ambiente permanente, residual e fenotípica, respectivamente. Para teor de proteína as variâncias estimadas variaram de 0,04 a 0,22; de 0,03 a 0,42 de 0,02 a 0,08 e de 0,10 a 0,62 para genética aditiva, ambiente permanente, residual e fenotípica, respectivamente. Para teor de gordura as variâncias estimadas variaram de 0,04 a 0,46; de 0,25 a 1,70; de 0,27 a 0,65 e de 0,61 a 2,44 para genética aditiva, ambiente permanente, residual e fenotípica, respectivamente. Já para teor de lactose as variâncias estimadas variaram de 0,02 a 0,03; de 0,02 a 0,13; de 0,02 a 0,04 e de 0,07 a 0,21 para genética aditiva, ambiente permanente, residual e fenotípica, respectivamente.

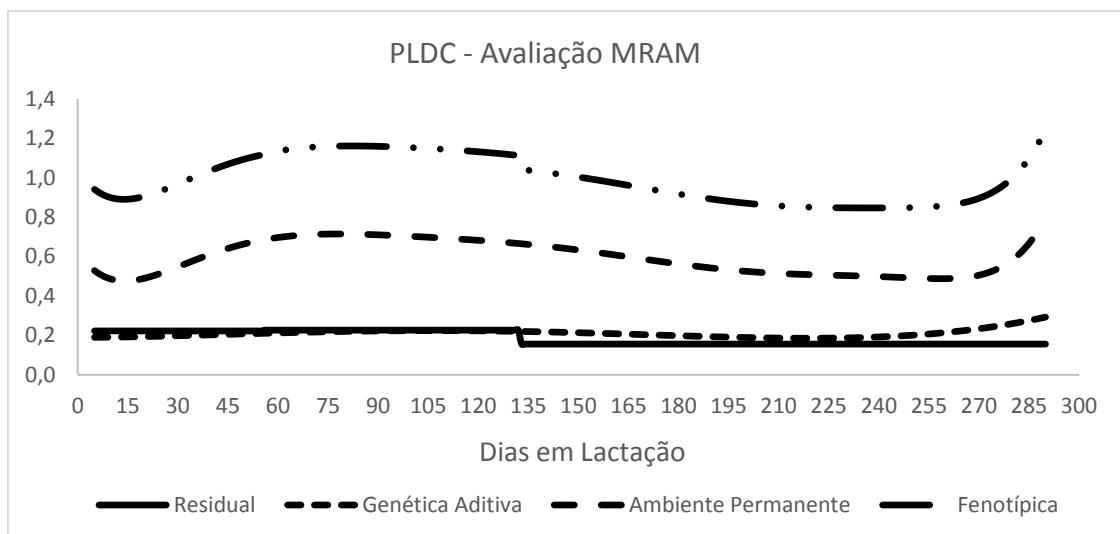


Figura 1: Variâncias genética aditiva, de ambiente permanente, fenotípica e residual para pldc obtida a partir da avaliação simultânea das quatro características (pldc, proteína, gordura e lactose) utilizando o MRAM.

As variâncias obtidas a partir do MRAM para pldc (Figura 1) apresentou o mesmo comportamento que o MRAU para a mesma característica. Em que a variância fenotípica apresentou crescimento a partir do 15º até o 75º dia de lactação e após a esse período os valores diminuem até o final da lactação. O mesmo se verifica para ambiente permanente, em que apresentou aumento até o 75º dia e após esse período volta a apresentar queda nos valores até o final

da lactação. Para genética aditiva é possível observar que conforme os dias em lactação iam aumentando os valores das estimativas também aumentavam, mesmo sendo um aumento discreto.

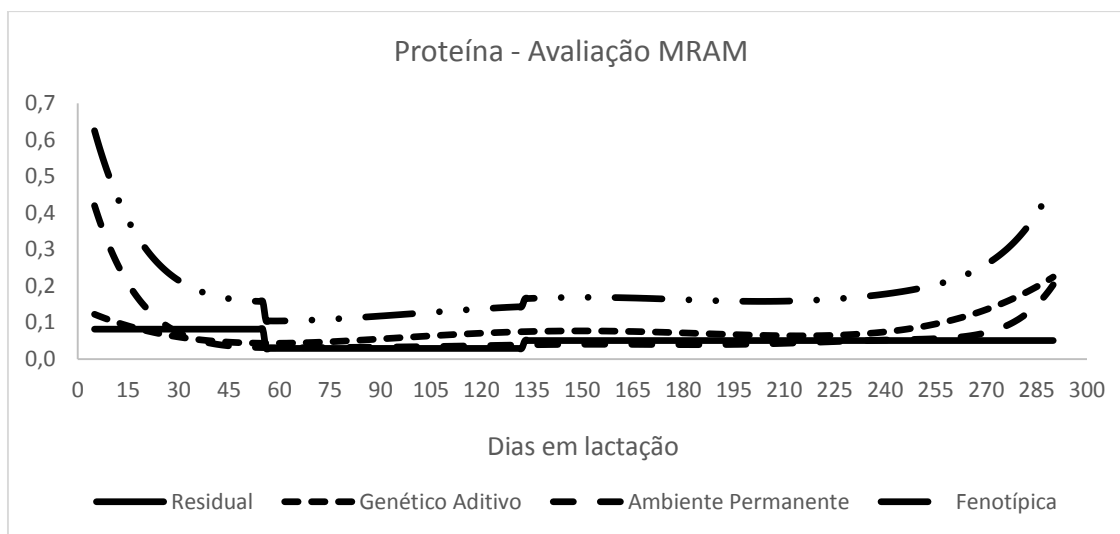


Figura 2: Variâncias genética aditiva, de ambiente permanente, fenotípica e residual para teor de proteína obtida a partir da avaliação simultânea das quatro características (pldc, proteína, gordura e lactose) utilizando o MRAM.

Também é possível observar semelhança no comportamento das variâncias no decorrer da lactação para teor de proteína utilizando MRAM (Figura 2) em relação ao MRAU. Crescimento no decorrer da lactação para variância genética é observado e que valores para variância fenotípica e de ambiente permanente apresentaram valores baixos a parti do 30º dia em lactação, com maiores valores nas extremidades da curva.

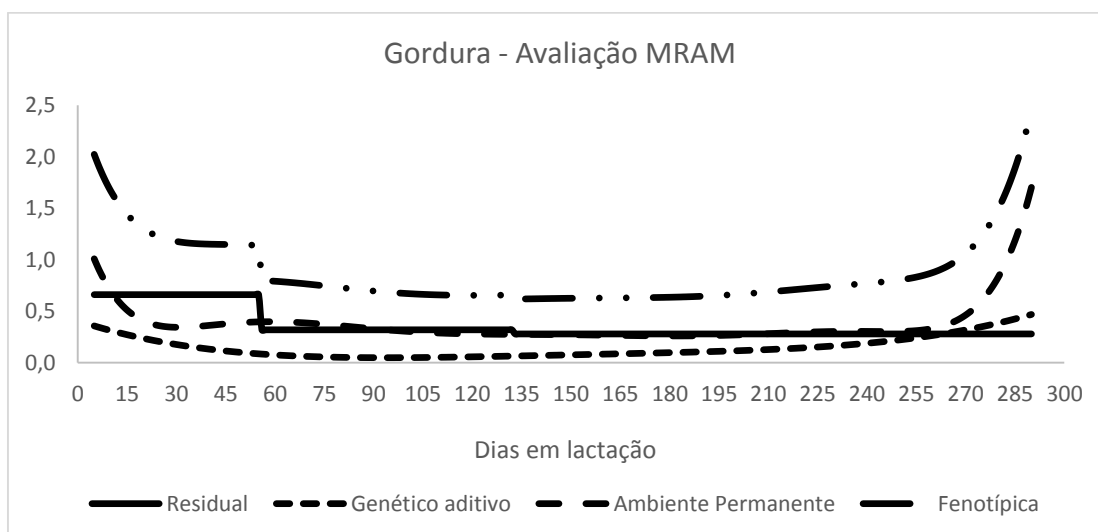


Figura 3: Variâncias genética aditiva, de ambiente permanente, fenotípica e residual para teor de gordura obtida a partir da avaliação simultânea das quatro características (pldc, proteína, gordura e lactose) utilizando o MRAM.

Para teor de gordura, o comportamento das variâncias para MRAM (Figura 3) apresentou algumas diferenças em relação ao MRAU. Uma das diferenças está na variância genética aditiva, em que apresentou decréscimo nos valores até o 75º dia e voltando a aumentar após o 120º até o final da lactação. Comportamento diferente observado no MRAU em que apresentou redução do início até o final dos dias em lactação. Já para a variância de ambiente permanente as extremidades foram o único momento que apresentou maiores valores e do 30º até o 270º foi o período em que menores valores foram observados com uma pequena variação no período compreendido entre o 45º e o 80º, o que não foi observado no MRAU. Para a variância fenotípica o mesmo comportamento foi observado para os dois tipos (MRAM e MRAU), sendo que os valores das estimativas foram diferentes.

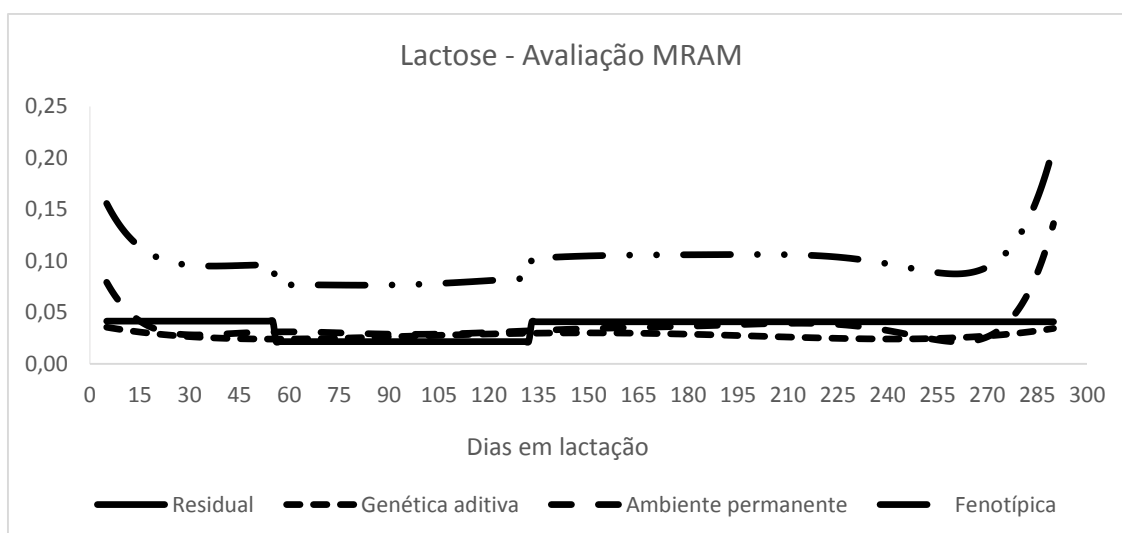


Figura 4: Variâncias genética aditiva, de ambiente permanente, fenotípica e residual para teor de lactose obtida a partir da avaliação simultânea das quatro características (pldc, proteína, gordura e lactose) utilizando o MRAM.

É possível observar que o comportamento das variâncias para teor de lactose utilizando MRAM (Figura 4) também foram semelhantes ao MRAU, em que valores baixos para a variância genética aditiva foram observados no meio da lactação e um pouco maiores no início e no final da lactação. Para a variância de ambiente permanente também apresentou os maiores valores nos extremos da lactação e no meio os valores foram baixos com pequenos aumentos nesse período. A variância fenotípica foi baixa no decorrer da lactação e alta no início e no final da lactação.

Herdabilidade

As estimativas de herdabilidade variaram de 0,19 a 0,26; de 0,20 a 0,53; de 0,07 a 0,31 e de 0,16 a 0,36 para as características pldc, proteína, gordura e lactose, respectivamente, obtidas a partir da utilização do MRAM (Figura 5).

É possível observar que a característica pldc apresentou herdabilidade de baixa a média magnitude e que maiores valores foram obtidos a partir do 165º dia de lactação, já no período em que o indivíduo apresenta o pico de produção a herdabilidade foram baixas. Para teor de proteína, o contrário é visto, em que maiores herdabilidade são observadas no período de alta produção de leite e após esse período a herdabilidade apresenta valores menores e constantes ao longo da lactação.

Para teor de gordura, as herdabilidade obtidas ao longo da lactação foram de baixa a média magnitude. No início até o 60º dia os valores diminuem e após a esse período as herdabilidades passam a aumentar com decorrer da lactação. Já para lactose os maiores valores foram obtidos no período compreendido entre o 60º e 135º dias de lactação, com o decorrer da lactação os valores diminuem e no final os valores voltaram a aumentar.

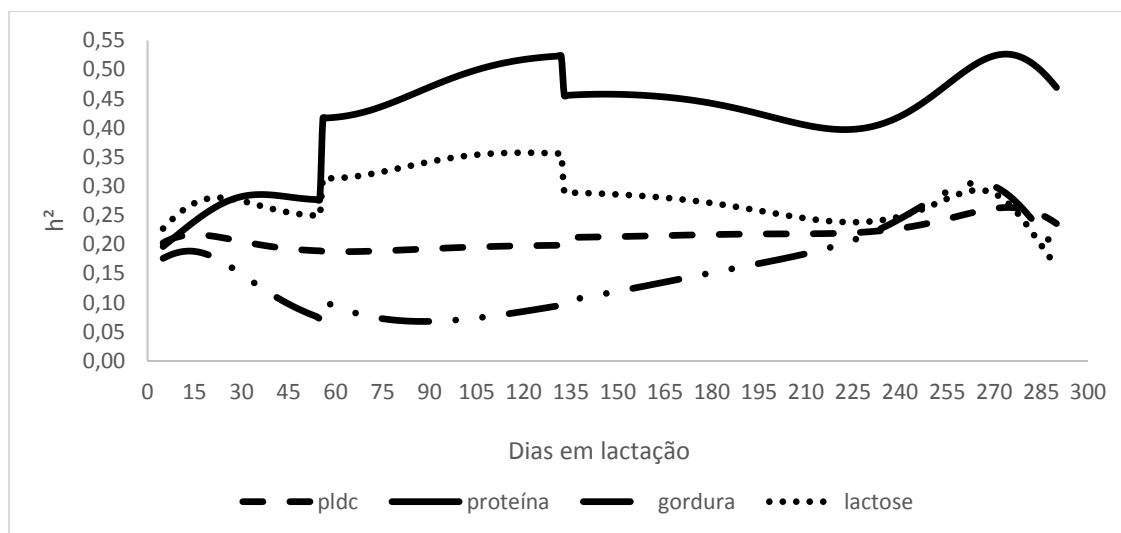


Figura 5: Estimativas de herdabilidade para PLDC, teor de proteína, gordura e lactose utilizando MRAM.

DISCUSSÃO

Cr terios de avalia o de modelos

No que diz respeito a escolha do modelo que melhor se ajustou aos dados, o F3A3P5H3   o recomendado para explicar as varia es das caracter sticas pldc, prote na, gordura e lactose, por apresentar melhor valor de BICm e o menor n mero de par metros, visto que modelos menos parametrizados s o desej veis. Em se tratando da ordem dos polin mios, Aspilcueta-Borquis et al. (2012) utilizaram modelos de regress o aleat ria em an lises individuais para rendimento de leite, gordura e prote na. Os autores recomendam que para todas as caracter sticas, polin mios de terceira ordem foram suficientes para modelar a vari ncia gen tica aditiva e recomenda que maiores ordens eram necess rias para ajustar o ambiente permanente (polin mios de quarta ordem para leite e prote na e quinta para a gordura).

Componentes de vari ncia e par metros gen ticos

As estimativas de vari ncias gen tica aditiva e de ambiente permanente obtidas para pldc, teor de prote na, gordura e lactose usando MRA unicaracter stico e multicaracter stico seguiram o mesmo padr o e valores de estimativas tamb m n o foram diferentes, somente para gordura que valores das estimativas foram um pouco maiores para o MRAU. Aspilcueta-Bosquis et al. (2013) utilizaram modelos de regress o aleat ria multicaracter sticos para produ o de leite, gordura e prote na em b falos, tamb m encontrou semelhan a entre os MRAU e MRAM com rela o a seguir o mesmo padr o, s  que estimativas maiores foram obtidas pela an lise MRAM, o que diferiu desse presente estudo. Os autores comentam que estes resultados podem provavelmente ser explicado porque, durante o processo de an lise de regress o todas as caracter sticas foram consideradas no MRAM e mais informa es s o avaliadas para a estima o dos componentes de (co)vari ncia.

Aspicueta-Borquis et al. (2010) relatam baixas estimativas para variância genética aditiva no início e no final da lactação e estimativas mais elevadas no terceiro mês de lactação. Resultado este que divergem com o presente estudo, onde foram encontradas maiores estimativas no final da lactação e baixos valores no início. Para a variância de ambiente permanente, os maiores valores foram observados no período de pico do indivíduo, o que mostra que é nesse período que a maiores influências do meio. Aspilcueta-Bosquis et al. (2013) e Sesana et al. (2010) observaram que maiores estimativas foram obtidas nos extremos da lactação e que em toda a lactação as estimativas foram quase constantes.

Com relação a variância de ambiente permanente para gordura e proteína, as estimativas apresentaram maiores valores nas extremidades da curva e durante a lactação os valores mantiveram baixos e constantes. Essa tendência também tem sido observada por Hammami et al. (2008) e Aspilcueta-Bosquis et al. (2013). Já para a variância genética aditiva apresentaram comportamento diferentes, em que estimativas para proteína aumentaram no decorrer da lactação e para gordura valores baixos foram encontrados no decorrer da lactação com maiores estimativas obtidas somente nas extremidades da curva. De Roos et al. (2004) utilizando regressão aleatória para rendimento de gordura e proteína realizados em bovino de leite, encontraram maiores estimativas no início e no final da lactação o que corroboram com o presente trabalho para a característica gordura.

Já as variâncias genética aditiva e de ambiente permanente para teor de lactose apresentaram o mesmo comportamento que a característica gordura. Comparações não foi possível ser feitas devido à escassez de estudo para lactose, mas é de conhecimento que a lactose tem um papel fundamental na síntese do leite.

Vale ressaltar que os valores das variâncias para teor gordura foram maiores que os de teores de proteína e lactose. Esse comportamento mostra que o teor de gordura é uma característica que apresenta maiores variações no decorrer da lactação de um indivíduo do que os outros constituintes.

A estimativa de herdabilidade para pldc e gordura apresentaram valores baixos no período compreendido entre 45° ao 120° e maiores valores nos extremos da curva. Estes resultados divergem de Aspilcueta-Bosquis et al. (2013), em que os autores encontraram comportamento inverso, com maiores valores no meio da lactação e menores valores nos extremos. Já para teor de proteína, as herdabilidades foram menores no início e maiores no período em que os indivíduos apresentam maiores produções de leite. Aspilcueta-Bosquis et al. (2010), utilizando um modelo dimensional finito relataram um comportamento semelhante de herdabilidade para gordura e proteína ao longo da lactação.

CONCLUSÕES

Modelos de regressão aleatória multicaracterístico apresentam maiores dificuldade de convergência do que os modelos de regressão aleatória unicaracterístico para cada característica.

O polinômio ortogonal de Legendre utilizado em análise de regressão aleatória multicaracterístico mais indicado para avaliação genética das características pldc, proteína, gordura e lactose foi o que considerou a ordem 3, 3 e 5 para a curva fixa e para os efeitos aleatórios, genético aditivo e de ambiente permanente, respectivamente, e ao menos 3 classes de variância residual.

REFERÊNCIA BIBLIOGRAFICA

- 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.
- ASPILCUETA-BORQUIS, R. B., R. C. SESANA, M. MUNOZ-BERROCAL, L. O. SENO, A. B. BIGNARDI, L. EL FARO, L. G. ALBUQUERQUE, G. M. CAMARGO, AND H. TONHATI. 2010. Genetic parameters for milk, fat and protein yields in Murrah buffaloes (*Bubalus bubalis* Artiodactyla, Bovidae). **Genet. Mol. Biol.** 33:71–77.

- ASPILCUETA-BORQUIS, R. B.; ARAUJO NETO, F. R.; BALDI, F.; HURTADO-LUGO, N.; DE CARMARGO, G. M. F.; MUÑOZ-BERROCAL, M.; and TONHATI, H. 2013. Multiple-trait random regression models for the estimation of genetic parameters for milk, fat, and protein yield in buffaloes **J. Dairy Sci.** 96:5923–5932.
- BREDA, F. C., L. G. ALBUQUERQUE, R. F. EUCLYDES, A. B. BIGNARDI, F. BALDI, R. A. TORRES, L. BARBOSA, AND H. TONHATI. 2010. Estimation of genetic parameters for milk yield in Murrah buffaloes by Bayesian inference. **J. Dairy Sci.** 93:784–791.
- DE ROOS, A. P. W., A. G. F. HARBERS, and G. DE JONG. 2004. Random herd curves in a test-day model for milk, fat and protein production of dairy cattle in the Netherlands. **J. Dairy Sci.** 87:2693-2701.
- HAMMAMI, H., B. REKIK, H. SOYEURT, A. BEN GARA, AND N. GENGLER. 2008. Genetic parameters for Tunisian Holsteins using a test-day random regression model. **J. Dairy Sci.** 91:2118–2126.
- JAMROZIK, J.; KISTEMAKER, G.J.; DEKKERS, J.C.M. et al. Comparison of possible covariates for use in a random regression model for analyses of test day yields. **Journal of Dairy Science**, v.80, n.10, p.2550-2556, 1997.
- 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.
- MUIR, B. L., G. KISTEMAKER, J. JAMROZIK, AND F. CANAVESI. 2007. Genetic parameters for a multiple-trait multiple-lactacao random regression test-day model in Italian Holsteins. **J. Dairy Sci.** 90:1564–1574.
- QUEIROGA, R. G. C.; BARRETO, T. M.; BISCONTINI, A. N. M.; MADRUGA, M. S.; RICARDO, A.; SCHULER, P. Influência do manejo do rebanho, das condições higiênicas da ordenha e da fase de lactação na composição química do leite de cabras Saanen. **R. Bras. Zootec**, v. 36, n. 2, p. 430-437, 2007.
- RAO, C.R. **Linear statistical inference and its applications**. 2.ed. New York: John Wilwy & Sons. p.417-420. 1973.
- SCHAEFFER, L. R., AND J. JAMROZIK. 2008. Random regression models: A longitudinal perspective. **J. Anim. Breed. Genet.** 125:145–146.
- SCHWARZ, G. Estimating the Dimension of a Model. **Annals of Statistics**, 1978, 6, p.461-464.

- SESANA, R. C., A. B. BIGANARDI, R. A. BORQUIS, L. EL FARO, F. BALDI, L. G. ALBUQUERQUE, AND H. TONHATI. 2010. Random regression models to estimate genetic parameters for test-day milk yield in Brazilian Murrah buffaloes. **J. Anim. Breed. Genet.** 127:369–376.
- 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.
- VAN DER WERF, J., AND L. SCHAEFFER. 1997. **Random regression in animal breeding.** Course notes. University of Guelph, Guelph, ON, Canada.
- WIGGANS, G. R., AND M. E. GODDARD. 1997. A computationally feasible test day model for genetic evaluation of yield traits in the United States. **J. Dairy Sci.** 80:1795–1800.

CONCLUSÕES GERAIS

Os modelos de regressão aleatória multicaracterístico e os unicaracterístico apresentaram estimativas semelhantes para os componentes de variância ao longo da curva de lactação para as características produção de leite no dia do controle, proteína e lactose, já a característica gordura foi a única que apresentou algumas diferenças nas estimativas das variâncias.

Mais estudos envolvendo modelos de regressão aleatória multicaracterístico para as mesmas características presente nesse estudo devem ser feitos, para que mais discussões possam ter com o objetivo de entender o comportamento das características e se à grande vantagem na utilização dos modelos de regressão aleatória multicaracterístico em vez dos modelos de regressão aleatória unicaracterístico.