

RESSALVA

Atendendo solicitação do(a)
autor(a), o texto completo desta tese
será disponibilizado somente a partir
de 03/03/2017.

UNIVERSIDADE ESTADUAL PAULISTA - UNESP
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL

**SELEÇÃO GENÔMICA PARA CARACTERÍSTICAS
LONGITUDINAIS DE BOVINOS DA RAÇA HOLANDESA**

Daiane Cristina Becker Scalez

Zootecnista

2016

UNIVERSIDADE ESTADUAL PAULISTA - UNESP
FACULDADE DE CIÊNCIAS AGRÁRIAS E VETERINÁRIAS
CÂMPUS DE JABOTICABAL

**SELEÇÃO GENÔMICA PARA CARACTERÍSTICAS
LONGITUDINAIS DE BOVINOS DA RAÇA HOLANDESA**

Daiane Cristina Becker Scalez

Orientadora: Profa. Dra. Sandra Aidar de Queiroz

Co-orientador: Dr. Haroldo Henrique de Rezende Neves

Tese apresentada à Faculdade de Ciências Agrárias e Veterinárias – UNESP, Câmpus de Jaboticabal, como parte das exigências para a obtenção do título de Doutor em Zootecnia.

DADOS CURRICULARES DO AUTOR

Daiane Cristina Becker Scalez nasceu na cidade de Rondonópolis - MT, em 13 de Agosto de 1988, filha de Marcos Antonio Scalez e Elenice Becker. Em 2006, ingressou no curso de Zootecnia na Universidade Federal do Mato Grosso – UFMT, Câmpus de Rondonópolis. Gradou-se em fevereiro de 2010 e, em março de 2010, ingressou no programa de Pós-Graduação em Ciência Animal (área de concentração: Produção e Melhoramento animal) da Universidade Federal do Mato Grosso, orientada pelo Prof. Dr. Fábio Luiz Buranelo Toral, obteve o título de Mestre em fevereiro de 2012. Em março de 2012, iniciou como aluna de doutorado do programa de Pós-Graduação em Zootecnia (área de concentração: Genética e Melhoramento dos Animais Domésticos) da Universidade Estadual “Júlio de Mesquita Filho”, Câmpus de Jaboticabal (UNESP – FCAV) tendo como orientadora a Prof. Dra Sandra Aidar de Queiroz. Em Abril de 2014, iniciou o estágio sanduíche no exterior, permanecendo doze meses na Universität Kassel na Alemanha, orientada pelo Prof. Dr. Sven König.

“Não há acaso, sina, destino, que possa limitar, impedir ou controlar
a firme resolução de uma alma determinada”

(Ella Wheller Wilcox)

Ao meu avô, Jaci Reinaldo Becker, que sempre me apoiou, acreditou e fez tudo que era possível para que eu realizasse meus sonhos.

DEDICO

À minha mãe, à minha vó e ao meu irmão que são a família que cuida de mim.

OFEREÇO

AGRADECIMENTOS

Agradeço...

A Deus por me iluminar a cada amanhecer.

À minha orientadora, Prof^a Dr^a Sandra Aidar de Queiroz, pela oportunidade de fazer o doutorado sob sua orientação, por me aceitar e me ajudar tanto durante esse período em Jaboticabal.

Ao meu co-orientador, Dr. Haroldo Henrique de Rezende Neves, pelo incentivo e pela imensa contribuição para que o trabalho fosse realizado.

Aos membros da banca, Prof. Henrique, Prof. Danísio, Dr^a Lenira e Dr^a Patrícia, pelas valiosas sugestões.

Ao Diercles, que me ajudou de todas as maneiras possíveis, não só com o suporte emocional, mas com valiosas contribuições para o trabalho. Muito obrigada por me ensinar a cada dia ser uma pessoa melhor e mais confiante, você é mais que um noivo, é um amigo e um companheiro de jornada.

Aos meus amigos da Pós-graduação, Cruz Elena, Soliris, Guilherme Venturini, Rafael Medeiros, Rafael Tonussi, Fabrícia, Danielly, Ana Cristina, Giovana, Henrique, Grazyella, Laiza, William, Tiago, Camila, Daniele, Daniel, Naudin, vocês tornaram o fardo mais leve e a caminhada muito mais divertida.

À minha família, que entendeu minhas escolhas e me deu todo o suporte para viver longe de casa, mesmo em momentos de perdas e muita tristeza.

À família do Diercles que me ajudou mais do que eu poderia imaginar que qualquer pessoa pudesse, me proporcionando um novo lar fora de casa.

Aos amigos de Pitangueiras, que me receberam tão bem todas as vezes que estive lá e me proporcionaram muitos finais de semana divertidos.

Ao Prof. Dr. Sven König, que me recebeu na Alemanha e me proporcionou adquirir mais conhecimento.

A todos da Universidade de Kassel na Alemanha, pela convivência e pelo aprendizado, especialmente à Tong Yin, que me ajudou e continua me ajudando todos os dias.

A Capes pela concessão da bolsa no início do doutorado.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela concessão da bolsa de doutorado no país (Processo n^o: 2013/ 01068-8) e pela bolsa no exterior (Processo n^o: 2013/23530-5).

Ao financiamento do BMBF e do FBF para o projeto colaborativo intitulado “KMU-innovativ-10: Kuh-L – cow calibration groups for the implementation of selection strategies based on high-density genotyping in dairy cattle”.

SUMÁRIO

	Página
1. INTRODUÇÃO	1
2. REVISÃO DE LITERATURA	3
2.1. Modelos de regressão aleatória (MRA)	3
2.2. Modelos não lineares – curvas de lactação	4
2.3. Seleção genômica	6
2.3.1. Metodologias de predição genômica e distribuição dos efeitos dos QTL ...	9
2.3.2. Validação da seleção genômica	13
3. OBJETIVOS	14
3.1. Objetivo geral	14
3.2. Objetivos específicos	14
4. MATERIAL E MÉTODOS	14
4.1. Descrição dos dados	14
4.2. Análise quantitativa	15
4.2.1. Edição dos dados – Modelos não lineares	15
4.2.2. Modelos não lineares	15
4.2.3. Edição dos dados para análises genômicas – modelos não lineares	16
4.2.4. Edição dos dados – Modelos de regressão aleatória	19
4.2.5. Modelos de regressão aleatória	20
4.3. Controle de qualidade dos dados genômicos – Modelos não lineares e modelos de regressão aleatória	22
4.4. Métodos de seleção genômica ampla (segundo estágio)	23
4.4.1. Comparação dos métodos	24
5. RESULTADOS E DISCUSSÃO	25
5.1. Ajuste das curvas individuais	25
5.2. Predição genômica – modelos não lineares	28
5.3. Modelo de regressão aleatória	32
5.4. Predição genômica – modelo de regressão aleatória	33
6. CONCLUSÕES	37
7. REFERÊNCIAS	38

SELEÇÃO GENÔMICA PARA CARACTERÍSTICAS LONGITUDINAIS DE BOVINOS DA RAÇA HOLANDESA

RESUMO - Características longitudinais são de grande importância para a produção animal, no entanto, quando se trata de seleção genômica vários métodos estão sendo avaliados para características pontuais, como produção de leite em idades específicas ou produção total. Dessa forma, a aplicação destes métodos a trajetórias completas do fenótipo de interesse ao longo do tempo pode ser uma ferramenta importante para a tomada de decisão do melhor momento para seleção dos animais. No presente estudo foram utilizados dados de produção de leite, porcentagem de gordura, porcentagem de proteína e escore de células somáticas de bovinos leiteiros da raça Holandesa em uma metodologia de dois estágios: i) os modelos não lineares de Wood (1967), Cobby & Le Du (1978) e Wilmink (1987) foram ajustados aos dados de cada característica; ii) as estimativas de parâmetros obtidas para as curvas individuais no modelo que proporcionou melhor ajuste aos dados (Wilmink) foram utilizadas como fenótipos nas análises genômicas. Além dos modelos não lineares, também foi utilizado um modelo de regressão aleatória (RRM), considerando polinômios de Legendre quadráticos para a modelagem dos efeitos genético aditivo e de ambiente permanente. No segundo estágio foram testados diferentes métodos para estimação dos valores genéticos genômicos (GEBV): RR-BLUP e LASSO, que consideraram como fenótipos as estimativas dos parâmetros obtidas anteriormente e os coeficientes de regressão estimados via RRM. A correspondência de fenótipos e GEBV foi verificada por meio de validação cruzada, utilizando o método k-fold. Para todas as características estudadas e todos os parâmetros estimados pelo modelo de Wilmink (1987), as acurácias da predição genômica foram baixas quando o RR-BLUP foi utilizado. Além disso, o parâmetro c apresentou acurácias próximas de zero para quase todas as características. As acurácias de predição obtidas para os coeficientes de regressão aleatória foram de magnitude moderada a alta ($>0,67$) demonstrando que o uso de coeficientes de regressão aleatória pode ser uma boa alternativa para predição genômica considerando dados longitudinais.

Palavras chave: LASSO Bayesiano, modelos de regressão aleatória, modelos não lineares, RR-BLUP

GENOMIC SELECTION FOR LONGITUDINAL TRAITS OF HOLSTEIN CATTLE

ABSTRACT - Longitudinal traits are of great importance for animal production, however, when it comes to genomic selection several methods are being evaluated for specific traits, such as milk yield in specific ages or total yield. Thus, applying these methods to complete trajectories of the phenotype of interest over time can be an important tool for helping to decide the best time to select the animals. In this study, milk yield, fat and protein percentage, and somatic cell score data of dairy cattle were used, in a two-stage procedure: i) where the nonlinear Wood (1967), Cobby & Le Du (1978) and Wilmink (1987) were adjusted; ii) the model that provided the best fit was the Wilmink model (1987), whose parameter estimates were used as phenotypes later. In addition to the non-linear models, it was also used a random regression model, considering quadratic Legendre polynomials to model the additive genetic and permanent environment effects. In the second stage different methods for estimation of genomic breeding values (GEBV) were applied: RR-BLUP and LASSO, which considered as phenotypes the estimated parameters previously obtained and the random regression coefficients. The correspondence between phenotypes and GEBV was verified by through cross validation using the k-fold method. For all traits, the genomic prediction accuracies of Wilmink's model parameters were low when RR-BLUP was used. Furthermore, the accuracies for parameter c showed values close to zero for most traits. The prediction accuracies for random regression coefficients were moderate to high, above 0.67, and demonstrated that the use of random regression coefficients can be a good alternative for genomic prediction using longitudinal data.

Keywords: Bayesian LASSO, random regression models, nonlinear models, RR-BLUP

1. INTRODUÇÃO

A União Europeia (UE) apresentou a maior produção mundial de leite em 2014, com 146,5 bilhões de litros produzidos, seguida pela Índia, Estados Unidos, China e Brasil. O bloco europeu também liderou a lista de maior exportador mundial de leite (USDA, 2015). Apesar de ser o maior produtor e exportador de leite do mundo, a atividade profissionalizada na UE é restrita a alguns países como, por exemplo, Alemanha, França, Dinamarca, Reino Unido e Espanha. Outros países, principalmente os do leste Europeu, têm pouca representatividade no total. Dentre os países citados, a Alemanha destaca-se, correspondendo a aproximadamente 20% da produção de leite da União Europeia, sendo o país de maior produção deste bloco (EUROSTAT, 2015).

Quando se trata de produtividade, a Alemanha ficou em terceiro lugar na UE, sendo os primeiros, Dinamarca e Holanda (EUROSTAT, 2015). A produção média na Alemanha é de 9.206 kg de leite por lactação, correspondendo à produção diária de 30,18 kg de leite. A média de produção das vacas de primeira lactação é de 8.000 kg, com média diária de 26,22 kg (DHV, 2015).

A raça Holandesa se destaca com o maior número de animais na Alemanha. O rebanho bovino alemão conta com 12,686 milhões cabeças, sendo que 4,267 milhões são vacas, dentre estas 2,123 milhões são da raça Holandesa, demonstrando a importância da raça para a produção de leite no país e na União Europeia (DHV, 2015).

As avaliações genéticas realizadas pela organização oficial de melhoramento da raça Holandesa na Alemanha (DHV, 2015) já estão levando em consideração a informação genômica para algumas características produtivas e para a fertilidade da fêmea. Dessa forma, os valores genéticos genômicos (GEBVs), que permitem a tomada de decisão baseada na combinação de informações fenotípicas e genotípicas, já estão sendo utilizados para a seleção de animais.

Neste contexto, vários métodos estatísticos vêm sendo utilizados para avaliação de características pontuais, como produção de leite em idades específicas ou produção total. Porém, isso ainda não vem sendo praticado para as

características longitudinais (i.e. aquelas tomadas ao longo do tempo, como produção de leite, de gordura e de proteína), situação em que as predições considerariam a trajetória de produção como um todo.

Inicialmente, essas características longitudinais eram avaliadas por meio de modelos não lineares, as chamadas curvas de lactação. Atualmente, a avaliação genética de medidas repetidas vem sendo realizada mais frequentemente por meio de *test-day models*, como modelos de regressão aleatória (RRM). Apesar da maior complexidade computacional, a adoção de métodos estatísticos que consideram a natureza longitudinal dos dados foi motivada por diferentes estudos que indicaram que a mesma proporcionaria maior precisão do que a avaliação da produção de leite acumulada aos 305 dias, além de permitir contemplar objetivos de seleção como persistência de produção.

Quando se trata de dados longitudinais, metodologias de avaliação genômica para curvas de crescimento têm sido propostas em alguns estudos (IBÁÑEZ-ESCRICHE; BLASCO, 2011; SILVA et al., 2013). Ibáñez-Escriche e Blasco (2011), considerando os três parâmetros da curva de crescimento de Gompertz e, utilizando simulação de dados, indicaram que a seleção genômica pode ser usada para alterar a forma da curva de crescimento de animais. No entanto, a obtenção de resposta em termos de alteração da curva completa (peso ao nascer, peso adulto e a forma da curva) não é fácil, devido às fortes correlações existentes entre todos os pesos corporais ao longo do crescimento do animal.

Em gado leiteiro, Koivula et al. (2012) utilizaram a metodologia de um único passo (*single-step*) aplicada a dados longitudinais de produção de leite, proteína, gordura e contagem de células somáticas e reportaram maior habilidade de predição deste método quando comparado a avaliações tradicionais em que a informação genômica não foi considerada.

Santos et al. (2014) propuseram o uso de coeficientes de regressão aleatória como fenótipos para a avaliação genômica de dados de produção de leite e consideraram essa abordagem viável, pois a habilidade de predição destes coeficientes foi satisfatória, sugerindo sua utilização para prever valores genéticos genômicos de animais jovens.

Desta forma, a aplicação de métodos de seleção genômica a trajetórias completas do fenótipo de interesse ao longo do tempo pode ser uma ferramenta importante para a tomada de decisão do melhor momento para seleção dos animais e deve ser estudada antes da sua aplicação.

6. CONCLUSÕES

O uso do modelo de Wilmink proporcionou acurácias relativamente baixas para pelo menos um parâmetro do modelo de Wilmink de todas as características. Vale ressaltar que o banco de dados utilizado para esta análise dispunha de poucas lactações completas, principalmente para as vacas genotipadas, ou seja, seriam necessários outros trabalhos com um banco de dados melhor estruturado para avaliar a adequação desta metodologia para predições genômicas.

Por outro lado, as acurácias de predição obtidas para os coeficientes de regressão aleatória foram de magnitude moderada a alta ($>0,67$), e com valores mais altos quando o método LASSO foi utilizado, demonstrando que o uso de coeficientes de regressão aleatória pode ser uma boa alternativa para predição genômica considerando dados longitudinais.

7. REFERÊNCIAS

- ARDLIE, K. G.; KRUGLYAK, L.; SEIELSTAD, M. Patterns of linkage disequilibrium in the human genome. **Nature reviews. Genetics**, v. 3, n. 4, p. 299–309, 2002.
- BIANCHINI SOBRINHO, E. Estudo da curva de lactação de vacas da raça Gir. Tese de Doutorado, USP, Ribeirão Preto, SP. 1984. 88 p.
- BRODY, S.; TURNER, C.W.; RAGSDALE, A.C. The rate of decline of milk secretion with the advance of the period lactation. **Journal of General Physiology**, v. 5, p. 441, 1923.
- CALUS, M. P. L. Genomic breeding value prediction: methods and procedures. **Animal**, v. 4, n.2, p. 157-164, 2010.
- COBBY, J.M. e LE DU, L.P. On fitting curves to lactation data. **Animal Production**, v. 26, p.127- 133,1978.
- CONGLETON, W.R.J.; EVERETT, R.W. Application of the Incomplete Gamma Function to predict cumulative milk production. **Journal of Dairy Science**, v. 63, p.109-19, 1980.
- DEKKERS, J. C. M. Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. **Journal of Animal Science**, v. 82 E-Suppl, p. E313–328, 2004.
- DE LOS CAMPOS, G.; NAYA, H.; GIANOLA, D.; CROSSA, J.; LEGARRA, A.; MANFREDI, E.; WEIGEL, K.; COTES, J. M. Predicting quantitative traits with regression models for dense molecular markers and pedigree. **Genetics**, v. 182, p. 375–385, 2009.

DAETWYLER, H. D.; PONG-WONG, R.; VILLANUEVA, B.; WOOLLIAMS, J. A. The impact of genetic architecture on genome-wide evaluation methods. **Genetics**, v. 185, p. 1021-1031, 2010.

Deutscher Holstein Verband e.V. – DHV. Statistic – Facts and figures 2014. Disponível em: <http://www.holstein-dhv.de/facts_and_figures.html>. Acesso em: 29 dez. 2015.

DRAPER, N.M.; SMITH, H. **Applied Regression Analysis**, 2.ed., New York, Wiley, 1981.

DURÃES, M. C.; TEIXEIRA, N. M.; FREITAS, A. F. Curvas de lactação de vacas da raça Holandesa mantidas em confinamento total. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.43, n.5, p.447-458, 1991.

ERBE, M.; PIMENTEL, E.C.G.; SHARIFI, A.R.; H. SIMIANER. Assessment of Cross-validation Strategies for Genomic Prediction in Cattle. In: 9th World Congress on Genetics Applied to Livestock Production, 2010, Leipzig. **Proceedings** of 9th World Congress on Genetics Applied to Livestock Production, 2010.

EUROSTAT Statistics Explained. Milk and milk products statistics. 2015. Disponível em: <http://ec.europa.eu/eurostat/statistics-explained/index.php/Milk_and_milk_product_statistics>. Acesso em: 30 dez. 2015.

FALCONER, D.S.; MACKAY, T.F.C. **Introduction to quantitative genetics**. 4.ed. Edinburgh : Longman Group Limited, 1996. 464p.

GAINES, W.L. Persistence of lactation in dairy cows. Agriculture Experiment Station, 1927. (Bulletin, 288).

- GARRICK, D. J.; TAYLOR, J. F.; FERNANDO, R. L. Deregressing estimated breeding values and weighting information for genomic regression analyses. **Genetics, Selection, Evolution**. p. 41-55. 2009.
- GIANOLA, D.; DE LOS CAMPOS, G.; HILL, W. G.; MANFREDI, E.; FERNANDO, R. Additive genetic variability and the Bayesian alphabet. **Genetics**, v. 183, p. 347-363, 2009.
- GODDARD, M. E.; HAYES, B. J. Genomic Selection. **Journal of Animal Breeding and Genetics**, v. 124, p. 323-330, 2007.
- GRISART, B.; FARNIR, F.; KARIM, L.; CAMBISANO, N.; KIM, J. J.; KVASZ, A.; MNI, M.; SIMON, P.; FRÈRE, J. M.; COPPIETERS, W.; GEORGES, M. Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition. **Proceedings of the National Academy of Sciences. USA**, v. 101, p. 2308-2403, 2004.
- GUO, G.; LUND, M. S.; ZHANG, Y.; SU, G. Comparison between genomic predictions using daughter yield deviation and conventional estimated breeding value as response variables. **Journal of Animal Breeding and Genetics**, v. 127, n.6, p.423–432. 2010.
- HABIER, D.; FERNANDO, R. L.; DEKKERS, J. C. M. The impact of genetic relationship on genome-assisted breeding values. **Genetics**, v. 117, p. 2389-2397, 2007.
- HABIER, D.; FERNANDO, R. L.; KIZILKAYA, K.; GARRICK, D. J. Extension of the Bayesian alphabet for genomic selection. **BMC Bioinformatics**, v. 12, n. 186. 2011.

- HAYES, B. J.; BOWMAN, P. J.; CHAMBERLAIN, A. J., GODDARD, M. E. Invited review: Genomic selection in dairy cattle: Progress and challenges. **Journal of Dairy Science**, v. 92, p. 433-443, 2009.
- HENDERSON, C. R. A simple method for computing the inverse of a numerator relationship matrix. used in prediction of breeding values. **Biometrics**, v. 32, p. 69-83, 1976.
- HENDERSON, C. R. **Applications of Linear Models in Animal Breeding**. University of Guelph, 1984.
- HENDERSON Jr., C. R. Analysis of covariance in the mixed model: higher-level, nonhomogeneous, and random regressions. **Biometrics**, v.38, p.623-640, 1982.
- HENSHALL, J. M.; GODDARD, M. E. Comparison of marker assisted selection using mixed model (BLUP) and mixed model with a test for the QTL. **Proceedings of 12th Conference Association for the Advancement of Animal Breeding and Genetics**, v. 12, p. 217–221, 1997.
- HUTCHISON, J. L.; COLE, J. B.; BICKHART, D. M. Short communication: Use of young bulls in the United States. **Journal of Dairy Science**, v. 97, p. 3213–3220, 2014.
- IBÁÑEZ-ESCRICHE, N.; BLASCO, A. Modifying growth curve parameters by multitrait genomic selection. **Journal of Animal Science**, v. 89, p. 661–668, 2011.
- KIRKPATRICK, M.; HECKMAN, N. A quantitative genetic model for growth, shape and other infinite-dimensional characters. **Journal of Mathematical Biology**, v.27, p.429-450, 1989.

- KIRKPATRICK, M.; LOFSVOLD, D.; BULMER, M. Analysis of the inheritance, selection and evolution of growth trajectories. **Genetics**, v.124, p.979-993, 1990.
- KOIVULA, M.; STRANDÉN, I.; PÖSÖ, J.; AAMAND, G. P.; MÄNTYSAARI, E. A. Single Step Genomic Evaluations for the Nordic Red Dairy Cattle Test Day Data. Interbull Bulletin n°. 46. Cork, Ireland, 2012.
- LANDE, R.; THOMPSON, R. Efficiency of marker-assisted selection in the improvement of quantitative traits. **Genetics**, v. 124, p. 743–756, 1990.
- LEGARRA, A.; RICARD A.; FILANGI O. **GS3 Genomic Selection — Gibbs Sampling — Gauss Seidel (and BayesCπ)**. Disponível em: <http://snp.toulouse.inra.fr/~alegarra/manualgs3_last.pdf>. Acesso em: 20 jul. 2015.
- LEGARRA, A.; ROBERT-GRANIE, C.; MANFREDI, E.; ELSEN, J-M. Performance of genomic selection in mice. **Genetics**, v. 180, p. 611–618, 2008.
- LEGARRA, A.; ROBERT-GRANIE, C.; CROISEAU, P.; GUILLAUME, F.; FRITZ, S. Improved Lasso for genomic selection. **Genetics Research**, v. 93, p. 77–87, 2011.
- LEGARRA, A. **Bases for Genomic Prediction**. Version 0.7.1. 77p. 2014.
- LIU, Z.; REINHARDT, F.; REENTS, R. Estimating Parameters of a Random Regression Test Day Model for First Three Lactation Milk Production Traits Using the Covariance Function Approach. Interbull Bulletin 01/2000; 25.
- LUND, M. S.; SAHANA, G.; KONING, D.; SU, G.; CARLBORG, O. Comparison of analyses of the QTLMAS XII common dataset. I: Genomic selection. **BMC Proceedings**, v.3(Suppl 1):S1, 2009.

- MEUWISSEN, T. H. E.; HAYES, B. J.; GODDARD, M. E. Prediction of total genetic value using genome-wide dense marker maps. **Genetics**, v. 157, p. 1819–1829, 2001.
- MEYER, K.; HILL, W. G. Estimation of genetic and phenotypic covariance functions for longitudinal or 'repeated' records by restricted maximum likelihood. **Livestock Production Science**, v.47, p.185-200, 1997.
- MEYER, K. Estimating covariance functions for longitudinal data using random regression model. **Genetics Selection Evolution**, v.30, p.221-240, 1998.
- MEYER, K. Scope for random regression model in genetic evaluation of beef cattle for growth. **Livestock Production Science**, v.86, p.69-83, 2004.
- MISZTAL, I.; TSURUTA, S.; STRABEL, T.; AUVRAY, B.; DRUET, T.; LEE, D. BLUPF90 and related programs (BGF90). In: 7th World Congress on Genetics Applied to Livestock Production, Montpellier. **Proceedings** of 7th World Congress on Genetics Applied to Livestock Production, Communication No. 28–07. INRA, Castanet-Tolosan, France, 2002.
- MOSER, G.; TIER, B.; CRUMP, R. E.; KHATKAR, M. S.; RAADSMA, H. W. A comparison of five methods to predict genomic breeding values of dairy bulls from genome-wide SNP markers. **Genetics Selection Evolution**, v. 41, p. 56, 2009.
- MRODE, R. A. Linear models for the prediction of animal breeding values. 2.ed. Wallingford: CABI Publishing, 2005. 344p.
- NEVES, H. H. R.; CARVALHEIRO, R.; O'BRIEN, A. M. P.; UTSUNOMIYA, Y. T.; CARMO, A. S.; SCHENKEL, F. S.; SÖLKNER, J.; MCEWAN, J.C.; VAN TASSEL, C. P.; COLE, J. B.; SILVA, M. V. G. B.; QUEIROZ, S. A.; SONSTERGARD, T. S.; GARCIA, J. F. Accuracy of genomic predictions in *Bos indicus* (Nellore) cattle. **Genetics Selection Evolution**, v. 46, n. 17, 2014.

PRYCE, J. E.; DAETWYLER, H. D. Designing dairy cattle breeding schemes under genomic selection: a review of international research. **Animal Production Science**, v. 52, p. 107-114, 2012.

PURCELL, S.; NEALE, B.; TODD-BROWN, K.; THOMAS, L.; FERREIRA, M. A.; BENDER, D.; MALLER, J.; SKLAR, P.; DE BAKKER, P. I.; DALY, M. J.; SHAM, P. C. PLINK: a tool set for whole-genome association and population-based linkage analyses. **American Journal of Human Genetics**, v. 81, n. 3, p. 559-575, 2007.

RESENDE, M. D. V. 2008 Genômica quantitativa e seleção no melhoramento de plantas perenes e animais. Colombo: Embrapa Florestas, 330 p.

RESENDE, M. D. V.; RESENDE Jr., M. F. R.; AGUIAR, A. M.; ABAD, J. I. M.; MISSIAGGIA, A. A.; SANSALONI, C.; PETROLI, C.; GRATTAPAGLIA, D. 2010 Computação da Seleção Genômica Ampla (GWS). Colombo: Embrapa Florestas, 79 p.

ROSO, V.M.; SCHENKEL, F.S. AMC – A computer program to assess the degree of connectedness among contemporary groups. In: 8th World Congress of Genetics Applied to Livestock Production, 2006, Belo Horizonte. **Proceedings** of 8th World Congress of Genetics Applied to Livestock Production, 2006. p.26. (Communication, 27).

SANTOS, D. J.; BOISON, S. A.; UTSUNOMYA, A. T.; PEIXOTO, M. G. C. D.; TONHATI, H.; SÖLKNER, J.; SILVA, M. V. An approach to genomic analysis of longitudinal data using random regression. In: 10th World Congress of Genetics Applied to Livestock Production, 2014, Vancouver, **Proceedings** of 10th World Congress of Genetics Applied to Livestock Production, 2014.

SAS User's Guide: Statistics. SAS Institute Inc., Cary, North Caroline, 2013.

- SEGELKE, D.; CHEN, J.; LIU, Z.; REINHARDT, F.; THALLER, G.; REENTS, R.
Reliability of genomic prediction for German Holsteins using imputed genotypes from low-density chips. **Journal of Dairy Science**, v.95, p. 5403-5411, 2012.
- SHOOK, G. E. Approaches to summarizing somatic cell count which improve interpretability. In: ANNUAL MEETING NATIONAL MASTITIS COUNCIL, 2., 1982, Arlington. **Proceedings...** Arlington: 1982. p.150.
- SILVA, K. M.; GUIMARÃES, S. E. F.; LOPES, P. S., NASCIMENTO, C. S.; LOPES, M. S.; AMARAL, M. M. D. C. Mapeamento de locos de características quantitativas para desempenho no cromossomo 4 de suínos. **Revista Brasileira de Zootecnia**, v. 38, p. 474-479, 2009.
- SILVA, F. F.; RESENDE, M. D. V.; ROCHA, G. S., DUARTE, D. A.; LOPES, P. S.; BRUSTOLINI, O. J.; THUS, S.; VIANA, J. M.; GUIMARÃES, S. E. Genomic growth curves of an outbred pig population. **Genetics and Molecular Biology**, v. 36, n. 4, p. 520-527, 2013.
- SOLBERG, T. R.; SONESSON, A. K.; WOOLLIAMS, J. A.; MEUWISSEN, T. H. Genomic selection using different marker types and densities. **Journal of Dairy Science**, v. 86, 2447–2454, 2008.
- SWALVE, H. H. Theoretical basis and computational methods for different test-day genetic evaluation methods. **Journal of Dairy Science**, v. 83, p. 1115-1124, 2000.
- TIBSHIRANI, R. Regression shrinkage and selection via the lasso. **Journal of the Royal Statistical Society, Series B**, v. 58, p. 267-288, 1996.
- UNITED STATES DEPARTMENT OF AGRICULTURE - USDA. **Dairy: World Markets and Trade**. Foreign Agricultural Service. Dec. 2015. 21p.

VAN DOORMAAL, B. Impact of Genomics on Genetic Selection and Gain. Canadian Dairy Network, p. 1–3, 2014.

WILMINK, J.B.M. Adjustment of test day milk, fat and protein yields for age, season and stage of lactation. **Livestock Production Science**, v.16, p.335-348, 1987.

WOOD, P.D.P. Algebraic model of the lactation curve in cattle. **Nature**, v.216, p.164-165, 1967.

APÊNDICE

Tabela 1a. Número de lactações e frequência de lactações de vacas da raça Holandesa, ajustados pelos modelos de Wood (1967), Cobby e Le Du (1978) e Wilmink (1987), de acordo com o R^2_A , para produção de leite

R^2_A	Wood		Cobby e Le Du		Wilmink	
	Número de lactações	Frequência (%)	Número de lactações	Frequência (%)	Número de lactações	Frequência (%)
$R^2_A \geq 0,50$	2.202	46,70	2077	44,05	2209	46,85
$R^2_A \geq 0,70$	1.354	28,72	1241	26,32	1336	28,34
$R^2_A \geq 0,80$	840	17,82	750	15,91	842	17,86
$R^2_A \geq 0,90$	270	5,73	270	5,73	284	6,02

Tabela 2a. Média, desvio-padrão (SD), coeficiente de variação (CV), valores mínimos e máximos das estimativas de parâmetros para as lactações individuais, ajustadas pelos modelos de Wood (1967), Cobby e Le Du (1978) e Wilmink (1987).

Parâmetro	Média	SD	CV (%)	Máximo	Mínimo
Wood 4.720 vacas					
a	31,9318	6,3371	19,8458	75,0491	10,6732
b	0,2537	0,3196	125,9707	2,7424	-2,8093
c	0,0838	0,0884	105,4960	0,9503	-0,7248
Cobby e Le Du 4.672 vacas					
a	36,1116	24,5444	67,9683	1.624,55	14,8063
b	0,0348	0,0631	181,3153	3,6550	-0,1151
c	1,5261E71	1,0214E73	6.692,99	6,98E74	0,0028
Wilmink 4.720 vacas					
a	35,6471	6,3594	17,8398	66,0422	6,0655
b	-0,0335	0,0325	-96,9222	0,1729	-0,2663
c	-12,8066	22,3124	-161,6066	251,8145	-260,4067

Tabela 3a. Média, desvio-padrão (SD), coeficiente de variação (CV), valores mínimos e máximos das estimativas de parâmetros para as lactações individuais, ajustadas pelos modelos de Wood (1967), Cobby e Le Du (1978) e Wilmink (1987)

Parâmetro	Média	SD	CV (%)	Mínimo	Máximo
Wood					
PP (4.313)	33,1034	5,2287	15,7951	9,4288	54,1491
TP (4.715)	46,8376	2620,59	5.595,05	-173.958,60	29224,89
Cobby e Le Du					
PP (4.156)	43,0503	25,8415	60,0262	2,5729E-73	221,5192
TP (4.156)	8,6909	16,1743	186,1066	0,2067	179,5871
Wilmink					
PP (3.865)	33,4363	5,1116	15,2877	16,4573	52,3665
TP (3.865)	62,2730	21,2221	34,0792	-66,5722	191,9300

PP: pico de produção (kg); TP: tempo para o pico de produção (dias).

Tabela 4a. Número de lactações e frequência de lactações de vacas Holandesas para porcentagem de gordura no leite, ajustados pelos modelos de Wood (1967), Cobby e Le Du (1978) e Wilmink (1987), de acordo com o R^2_A

R^2_A	Wood		Cobby e Le Du		Wilmink	
	Número de lactações	Frequência (%)	Número de lactações	Frequência (%)	Número de lactações	Frequência (%)
$R^2_A \geq 0,50$	1.990	42,10	1	0	1.924	40,70
$R^2_A \geq 0,70$	1.180	24,96	0	0	1.103	23,33
$R^2_A \geq 0,80$	728	15,40	0	0	654	13,84
$R^2_A \geq 0,90$	276	5,84	0	0	236	4,99

Tabela 5a. Média, desvio-padrão (SD), coeficiente de variação (CV), valores mínimos e máximos das estimativas de parâmetros para as lactações individuais, ajustadas pelos modelos de Wood (1967), Cobby e Le Du (1978) e Wilmink (1987), para porcentagem de gordura

Parâmetro	Média	SD	CV (%)	Máximo	Mínimo
Wood 4.727 vacas					
a	8,7348	33,8154	387,1326	1990,98	0,0488
b	-0,1603	0,2118	-132,0935	1,3761	-1,8721
c	-0,0020	0,0022	-110,6056	0,0206	-0,0189
Cobby e Le Du 3.631 vacas					
a	3623583199	2140346699	59,0671328	269467206199	431580,70
	8	6			
b	3623,53	2.140,33	59,0676427	26946,54	0,0010
c	1E-7	0	0	1E-7	1E-7
Wilmink 4.727 vacas					
a	3,2213	0,6482	20,1215	7,2859	-1,4940
b	0,0038	0,0040	104,7239	0,0393	-0,0300
C	2,1078	2,9942	142,0510	29,4456	-19,3320

Tabela 6a. Média, desvio-padrão (SD), coeficiente de variação (CV), valores mínimos e máximos das estimativas de parâmetros para as lactações individuais, ajustadas pelos modelos de Wood (1967), Cobby e Le Du (1978) e Wilmink (1987), para porcentagem de gordura

Parâmetro	Média	SD	CV (%)	Mínimo	Máximo
Wood					
PP (4.302)	3,5585	0,5023	14,1147	1,8653	7,7203
TP (4.727)	33,6580	5329,68	15834,80	-320074,91	168134,36
Cobby e Le Du					
PP (3.631)	43907,81	1318114,34	3002,00	62,8062	41371450,42
TP (3.631)	0,0044	0,1318	3002,00	6,2806E-6	4,1371
Wilmink					
PP (3.929)	3,4858	0,4610	13,2247	1,4426	5,6202
TP (3.929)	67,0482	22,0045	32,8189	-58,7653	180,2080

PP: pico de produção (kg); TP: tempo para o pico de produção (dias).

Tabela 7a. Número de lactações e frequência de lactações de vacas Holandesas para porcentagem de proteína no leite, ajustados pelos modelos de Wood (1967), Cobby e Le Du (1978) e Wilmink (1987), de acordo com o R^2_A

R^2_A	Wood		Cobby e Le Du		Wilmink	
	Número de lactações	Frequência (%)	Número de lactações	Frequência (%)	Número de lactações	Frequência (%)
$R^2_A \geq 0,50$	3149	66,50	2198	46,42	3238	68,38
$R^2_A \geq 0,70$	2237	47,24	1379	29,12	2341	49,44
$R^2_A \geq 0,80$	1555	32,84	870	18,37	1614	34,09
$R^2_A \geq 0,90$	631	13,33	331	6,99	679	14,34

Tabela 8a. Média, desvio-padrão (SD), coeficiente de variação (CV), valores mínimos e máximos das estimativas de parâmetros para as lactações individuais, ajustadas pelos modelos de Wood (1967), Cobby e Le Du (1978) e Wilmink (1987), para porcentagem de proteína no leite

Parâmetro	Média	SD	CV (%)	Máximo	Mínimo
Wood 4.735 vacas					
a	3,3514	1,1471	34,2268	14,6080	0,1879
b	-0,0148	0,1084	-731,2554	0,7409	-0,4333
c	-0,0009	0,0011	-114,7322	0,0061	-0,0062
Cobby e Le Du 4.719 vacas					
a	3,0595	0,2530	8,2690	4,4829	2,1771
b	-0,0021	0,0015	-69,2989	0,0084	-0,0105
c	1,7802	6,1911E71	3477,75	3,0631E73	0,0201
Wilmink 4.735 vacas					
a	2,9241	0,2857	9,7716	4,1855	1,7433
b	0,0029	0,0016	55,0681	0,0166	-0,0064
c	0,2422	1,2391	511,5974	7,0825	-12,6656

Tabela 9a. Média, desvio-padrão (SD), coeficiente de variação (CV), valores mínimos e máximos das estimativas de parâmetros para as lactações individuais, ajustadas pelos modelos de Wood (1967), Cobby e Le Du (1978) e Wilmink (1987), para porcentagem de proteína no leite

Parâmetro	Média	SD	CV (%)	Mínimo	Máximo
Wood					
PP (3.634)	3,1880	0,3349	10,5049	2,5401	6,8473
TP (4.735)	-55,6979	6744,95	-12109,88	-423038,85	102473,56
Cobby e Le Du					
PP (293)	28,7530	46,4148	161,4255	5,8145E-72	187,6391
TP (293)	26,9518	36,0620	133,8019	2,9687	178,1038
Wilmink					
PP (3.298)	3,0574	0,2116	6,9208	2,0859	3,8943
TP (3.298)	48,0051	22,2559	46,3616	-134,2824	167,3677

PP: pico de produção (kg); TP: tempo para o pico de produção (dias).

Tabela 10a. Número de lactações e frequência de lactações de vacas Holandesas para escore de células somáticas, ajustados pelos modelos de Wood (1967), Cobby e Le Du (1978) e Wilmink (1987), de acordo com o R^2_A

R^2_A	Wood		Cobby e Le Du		Wilmink	
	Número de lactações	Frequência (%)	Número de lactações	Frequência (%)	Número de lactações	Frequência (%)
$R^2_A \geq 0,50$	988	20,99	242	5,14	971	20,63
$R^2_A \geq 0,70$	545	11,58	100	2,12	509	10,82
$R^2_A \geq 0,80$	322	6,84	49	1,04	308	6,54
$R^2_A \geq 0,90$	136	2,89	17	0,003	122	2,59

Tabela 11a. Média, desvio-padrão (SD), coeficiente de variação (CV), valores mínimos e máximos das estimativas de parâmetros para as lactações individuais, ajustadas pelos modelos de Wood (1967), Cobby e Le Du (1978) e Wilmink (1987), para escore de células somáticas

Parâmetro	Média	SD	CV (%)	Máximo	Mínimo
Wood 4.706 vacas					
a	-1,6286E51	1,0824E53	-6645,76	5,1933E31	-7,4210E51
b	0,0241	2,5722	10688,18	26,4960	-41,0343
c	1502917250	1,0310E13	6860,03	7,0727E14	-0,3686
	73				
Cobby e Le Du 3.625 vacas					
a	8711,77	500210,11	5741,77	30056103,75	-531406,78
b	-0,0115	1,2971	-11259,73	10,7246	-76,5320
c	1,9241	7,6792E72	3991,00	4,1786	-0,2163
Wilmink 4.706 vacas					
a	1,9729	1,8801	95,2974	12,4652	-5,6946
b	0,0019	0,0100	532,8249	0,0769	-0,0780
c	1,8152	7,4479	410,3060	58,3704	-103,2317

Tabela 12a. Média, desvio-padrão (SD), coeficiente de variação (CV), valores mínimos e máximos das estimativas de parâmetros para as lactações individuais, ajustadas pelos modelos de Wood (1967), Cobby e Le Du (1978) e Wilmink (1987), para escore de células somáticas

Parâmetro	Média	SD	CV (%)	Mínimo	Máximo
Wood (
PP (4.151)	2,4387	5,9698	244,7916	-0,9083	374,2324
TP (4.706)	-54,0433	9765,37	-18069,52	-603561,18	189684,41
Cobby e Le Du					
PP (2.647)	35,3678	65,5408	185,3119	-987,8765	617,1762
TP (2.647)	20,8228	31,6270	151,8864	-29,9353	177,2752
Wilmink					
PP (3.448)	2,0379	1,4244	69,8973	-0,9337	8,4712
TP (3.448)	70,2569	27,6261	39,3215	-99,5512	223,5604

PP: pico de produção (kg); TP: tempo para o pico de produção (dias).