

DIEGO HELCIAS CAVALCANTE

**FUNÇÕES DE COVARIÂNCIAS NA MODELAGEM DO CRESCIMENTO DE
BOVINOS NELORE MOCHO CRIADOS NA REGIÃO NORTE DO BRASIL**

TERESINA - PI

2017

DIEGO HELCIAS CAVALCANTE

**FUNÇÕES DE COVARIÂNCIAS NA MODELAGEM DO CRESCIMENTO DE
BOVINOS NELORE MOCHO CRIADOS NA REGIÃO NORTE DO BRASIL**

Defesa de Tese apresentada ao Programa de Pós-Graduação em Ciência Animal da Universidade Federal do Piauí (UFPI), como requisito a obtenção do grau de Doutor em Ciência Animal.
Área de Concentração: Produção Animal

Orientador: Prof. Dr. Severino Cavalcante Sousa Júnior

TERESINA – PI

2017

FICHA CATALOGRÁFICA
Universidade Federal do Piauí
Biblioteca Setorial do Centro de Ciências Agrárias
Serviço de Processamento Técnico

C377f Cavalcante, Diego Helcias
Funções de covariâncias na modelagem do crescimento de
bovinos nelore mocho criados na região norte do Brasil / Diego
Helcias Cavalcante - 2017.
92 f.: il.

Tese (Doutorado em Ciência Animal) – Universidade Federal
do Piauí, Teresina, 2017.
Orientação: Prof. Dr. Severino Cavalcante Sousa Júnior

1.Gado Nelore 2. Dados longitudinais 3.Modelos de regressão
aleatória 4. Legendre 5.B-splines I.Título

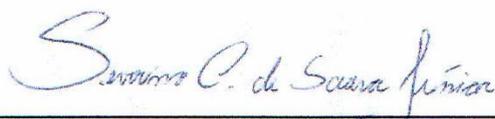
CDD 636. 291

FUNÇÕES DE COVARIÂNCIAS NA MODELAGEM DE CRESCIMENTO DE BOVINOS NELORE MOCHO CRIADOS NA REGIÃO NORTE DO BRASIL

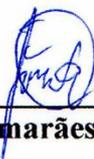
DIEGO HELCIAS CAVALCANTE

Tese Aprovada em: 23/03/2017

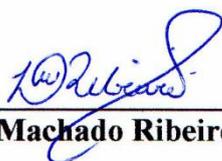
Banca Examinadora:



Prof. Dr. Severino Cavalcante de Sousa Junior (Presidente) / CMRV/UFPI



Prof. Dr. José Elivalto Guimarães Campelo (Interno) / DZO/CCA/UFPI



Profa. Dra. Danielle Maria Machado Ribeiro Azevedo (Interna) / EMBRAPA



Prof. Dr. Luciano Pinheiro da Silva (Externo) / UFC



Prof. Dr. Raimundo Martins Filho (Externo) / UFC

“Um caminho é só um caminho, e não há desrespeito a si ou aos outros em abandoná-lo, se é isto que o coração nos diz...
Examine cada caminho com muito cuidado e deliberação.
Tente-o muitas vezes, tanto quanto julgar necessário.
Só então pergunte a você mesmo, sozinho, uma coisa...
Este caminho tem coração?
Se tem, o caminho é bom,
se não tem, ele não lhe serve.
Um caminho é só um caminho”.

Carlos Castañeda

Aos meus pais, José Ubiratan Cavalcante e Lúcia Maranhão Helcias Cavalcante, exemplos de trabalho e honestidade.

Aos meus irmãos, Thiago Helcias Cavalcante, Rayssa Helcias Cavalcante e Glauber Helcias Cavalcante (in memoriam).

A todos que contribuíram para a realização desse trabalho.

Dedico.

Agradecimentos

À Universidade Federal do Piauí – UFPI e ao Programa de Pós-Graduação em Ciência Animal do Centro de Ciências Agrárias da Universidade Federal do Piauí, pela realização do curso;

Aos Docentes do Curso de Doutorado em Ciência Animal do Centro de Ciências Agrária da Universidade Federal do Piauí, em especial: o Prof. Dr. Severino Cavalcante de Sousa Júnior, Dr. José Elivalto Guimarães Campelo, Dr. José Lindenberg Rocha Sarmento, Dra. Danielle Maria Machado Azevedo pelos ensinamentos nas disciplinas, na vida e pela amizade;

Aos professores Dr. Carlos Henrique Mendes Malhado da UESB – BA e Dr. Luciano Pinheiro da Silva da UFC – CE, ambos com contribuições fundamentais para realização deste trabalho;

Aos meus pais, José Ubiratan Cavalcante e Lúcia Maranhão Helcias Cavalcante, e, extendendo meus agradecimentos a todos os familiares: obrigado pelo apoio e palavras de incentivo!;

À minha namorada Ana Carolina Portela por alegrar meus dias e por estar sempre ao meu lado, nos momentos difíceis assim como nos mais felizes!;

Aos colegas de pós-graduação que eu convivi neste período de pesquisa: Agrícola, Samy, Raizza, Mabell, Snaylla, Aurino, Natanael, Laylson, Luciano, Ivamara, Tatiana, Vanessa, Débora, Bruna, João, Amauri, Weverton, Marcelo, Syllas e a todos os demais colegas do GEMA, pelos momentos de estudo, troca de conhecimento e descontração;

E, concluindo, agradeço a Deus por colocar essas pessoas na minha trajetória de vida.

Obrigado!

Sumário

Lista de Figuras	ix
Lista de Tabelas	xi
Resumo	xiii
Abstract	xiv
1 INTRODUÇÃO GERAL	15
2 REVISÃO DE LITERATURA	17
2.1 A bovinocultura de corte no Brasil	17
2.2 Modelagem do crescimento animal.....	19
2.3 Modelos de regressão aleatória e funções de covariância	20
2.4 Funções B-splines	23
3 REFERÊNCIAS BIBLIOGRÁFICAS DA INTRODUÇÃO GERAL E REVISÃO	25
4 CAPÍTULO I	28
Modelagem da curva fixa de crescimento de bovinos Nelore Mocho da região Norte do Brasil	28
Palavras-chave: curva fixa, dados longitudinais, modelos de regressão aleatória, Legendre, Nelore.	28
Introdução.....	30
Material e Métodos.....	31
Resultados e Discussão	35
Conclusões	40
Referências Bibliográficas	41
5 CAPÍTULO II	43
Modelos de regressão aleatória sob diferentes estruturas de variância residual para descrever o crescimento de bovinos Nelore Mocho	43
Introdução.....	45
Material e Métodos.....	45
Resultados e Discussão	49
Conclusões	54
Referências Bibliográficas	54
6 CAPÍTULO III	56
Funções de covariâncias sob polinômios de Legendre para modelagem do crescimento de bovinos da raça Nelore Mocho da região Norte do Brasil.....	56
Introdução.....	58
Material e Métodos.....	58
Resultados e Discussão	63
Conclusões	70
Referências Bibliográficas	70

7 CAPÍTULO IV	73
Funções de covariâncias sob polinômios B-spline para modelagem do crescimento de bovinos da raça Nelore Mocho da região Norte do Brasil	73
Introdução.....	75
Material e método	76
Resultados e discussão.....	81
Conclusão	89
Referências Bibliográficas	90
7 CONSIDERAÇÕES FINAIS	92

Lista de Figuras

Figura 1. Produção mundial de carne bovina, em porcentagem (USDA, 2016).	17
Figura 1.1. Número de registro e peso corporal médio por idade de bovinos Nelores provenientes da região norte do Brasil.	32
Figura 1.2. Dispersão dos pesos médios observados em função da idade e trajetória média populacional estimada (regressão fixa sob polinômios ortogonais da idade) por quadrados mínimos de acordo com a ordem de ajuste (K).	37
Figura 1.3. Estimativas de herdabilidade direta (h^2) e de componente de variância para ambiente permanente animal como proporção da variância fenotípica (c^2), obtidas por meio de modelos que empregaram funções de diferentes ordens para representar a trajetória média.	38
Figura 1.4. Estimativas de herdabilidade materna (m^2) e de componente de variância para ambiente permanente materno como proporção da variância fenotípica (q^2) obtidas por meio de modelos que empregaram funções de diferentes ordens para representar a trajetória média.	39
Figura 2.1. Número de registro e peso corporal médio por idade de bovinos Nelores provenientes da região Norte do Brasil.	47
Figura 2.2. Estimativas de variâncias residuais homogênea (CL1) e heterogêneas por meio das classes CL3, CL4 e CL6.	50
Figura 2.3. Estimativas de herdabilidade direta (h^2 - acima à esquerda) e materna (m^2 - acima à direita) e componentes de variância para ambiente permanente de animal (c^2 - abaixo à esquerda) e materno (q^2 - abaixo à direita) como proporção da variância fenotípica.....	51
Figura 2.4. Correlação de Spearman para o valor genético aditivo direto, obtido em diferentes idades, por meio do modelo composto por 6 classes de variância residual (CL6).....	53
Figura 3.1. Número de registro e peso corporal médio por idade de bovinos Nelores provenientes da região norte do Brasil.	60
Figura 3.2. Estimativas de componentes de variância genética aditiva direta e genética materna (acima à esquerda), de ambiente permanente de animal de ambiente permanente materno (acima à direita), metade da variância fenotípica (abaixo à esquerda) e variância residual (abaixo à direita).....	67
Figura 3.3. Estimativas de herdabilidade direta (h^2) e materna (m^2) e componentes de variância para ambiente permanente de animal (c^2) e materno (q^2) como proporção da variância fenotípica, nas 117 classes de idade, utilizando modelo de regressão aleatória.....	69

Figura 4.1. Número de registros e peso médio por idade de bovinos da raça Nelore Mocho criados na região Norte do Brasil.	77
Figura 4.2. Estimativas de componentes de variância aditiva genética direta (s^2a), genética materna (s^2m), ambiente permanente direto(s^2pe), ambiente permanente materno (s^2mpe), residual(s^2e) e fenotípica(s^2f), para os modelos Leg6333, bspL_6633, bspQ_5544, e bspC_6555	83
Figura 4.3. Estimativas de herdabilidade direta (h^2), herdabilidade materna (m^2), estimativa da variância de ambiente permanente de animal (c^2) e materno (q^2) como proporção da variância fenotípica total, obtidas por modelos de regressão aleatória para os modelos de regressão aleatória.	85
Figura 4.4. Correlação de Spearman para o valor genético aditivo direto, obtido em diferentes classes de idades, por meio do modelo bspC_6555.....	88

Lista de Tabelas

Tabela 1.1. Resumo da estrutura dos dados de bovinos Nelores provenientes da região Norte do Brasil utilizados no estudo.....	32
Tabela 1.2. Quadrado médio do erro de predição (QMR), desvio médio absoluto (DMA), coeficiente de determinação (R^2), percentagem quadrada de viés (PQV) para o ajuste da trajetória média da população por uma função polinomial com diferentes ordens (k)	36
Tabela 1.3. Correlação de Spearman entre os valores genéticos preditos para os pesos à desmama (acima da diagonal) e aos 660 dias de idade (abaixo da diagonal) por meio de modelos que empregaram funções de diferentes ordens para representar a trajetória média da po	40
Tabela 2.1.-Resumo da estrutura dos dados de bovinos Nelores provenientes da região Norte do Brasil utilizados no estudo	47
Tabela 2.2. Número de parâmetros (NP), Logaritmo da função de máxima verossimilhança (LogL), critério de informação de Akaike (AIC), critério de informação Bayesiano (BIC) e teste da razão de verossimilhança (TRV) estimados com os modelos que empregaram diferentes	49
Tabela 2.3. Valores genéticos preditos para os cinco melhores animais para pesos ao nascimento, aos 240 e 660 dias de idade considerando homogeneidade e heterogeneidade (CL6) de variância residual.	52
Tabela 3.1. Resumo da estrutura dos dados de bovinos Nelores provenientes da região norte do Brasil utilizados no estudo.....	60
Tabela 3.2. Análise descritiva dos dados de peso à idade referentes aos bovinos Nelore Mocho criados na região norte do Brasil.	63
Tabela 3.3. Logaritmo da função de máxima verossimilhança (Log L), critério de informação de Akaike (AIC), critério de informação Bayesiano (BIC), número de parâmetros estimados (NP), expressos como desvio absoluto do menor valor.....	64
Tabela 3.4. Estimativas de variâncias (diagonal), covariâncias (abaixo da diagonal) e correlações (acima da diagonal) entre os coeficientes de regressão aleatória e autovalores associados à matriz dos coeficientes para o modelo com $k= 6333$	66
Tabela 4.1-Resumo da estrutura dos dados de bovinos Nelore Mocho provenientes da região Norte do Brasil.	77
Tabela 4. 2. - Modelos, número de parâmetros (n) do modelo, posição dos nós (classe de idade) para os efeitos genéticos direto e materno, números de segmentos do efeito genético direto (s),	

critérios de informação de Akaike (AIC) e Bayesiano de Schwarz BIC), para os modelos de regressão aleatória. 81

Tabela 4.3. Estimativas de variâncias (diagonal), covariâncias (abaixo da diagonal) e correlações (acima da diagonal) entre os coeficientes de regressão aleatória e autovalores associados à matriz dos coeficientes para o modelo bspC_6555. 86

Resumo: Funções polinomiais da idade com diferentes ordens foram avaliadas na modelagem do crescimento de bovinos Nelore Mocho utilizando modelos de regressão aleatória (MRA). Utilizaram-se 15.148 registros de pesos de 3.115 bovinos da raça Nelore Mocho com idade entre 1 e 660 dias, criados na região Norte do Brasil e nascidos entre os anos de 1995 e 2010. Com base nos resultados apresentados nesta Tese, os polinômios ortogonais da idade mostraram-se eficientes para modelagem da curva média através de um polinômio de quarta ordem e para modelagem da variância residual, o modelo de regressão aleatória com 6 classes heterogências permitiu um melhor ajuste aos dados de crescimento dos animais em estudo. Posteriormente, para a modelagem da parte aleatória, foram consideradas funções de diferentes ordens para modelar as variâncias associadas aos efeitos genéticos aditivos direto e materno e aos efeitos de ambiente permanente do animal e da mãe. De acordo com o valor dos critérios AIC e BIC o modelo leg_6333 permitiu ajustar as mudanças das variâncias e covariâncias com a idade dos bovinos Nelore Mocho. No último capítulo desta Tese, comparou-se diferentes modelos de regressão aleatória ajustados por polinômios segmentados do tipo B-spline e pelo polinômio ortogonal de Legendre (leg_6333). Embora o modelo que utilizou os polinômios ortogonais de Legendre (leg_6333) tenha sido mais parcimonioso, o modelo ajustando funções B-spline cúbicas, (bspC_6555), com quatro nós e três segmentos para os efeitos genético aditivo direto e, três nós e dois segmentos para os efeitos genético materno, ambiente permanente de animal e ambiente permanente materno, foi capaz de descrever de forma adequada, e com maior velocidade de convergência, a estrutura de (co)variâncias do conjunto de dados de peso de bovinos Nelore Mocho criados na região Norte do Brasil.

Palavras-chave: gado nelore, dados longitudinais, modelos de regressão aleatória, Legendre, B-splines.

Abstract: Polynomial functions of age with different orders were evaluated in the modeling of growth of Nelore owl cattle using random regression models (MRA). A total of 15,148 weight records were collected from 3,115 Nelore Mocho cattle aged 1 to 660 days from the northern region of Brazil, born between 1995 and 2010. Based on the results presented in chapter 1 of this thesis, Orthogonal polynomials of age were efficient for modeling the mean curve through a fourth-order polynomial. In Chapter 2, the random regression model with 6 heterogeneous classes of residual variances allowed a better fit in the growth modeling of the animals under study. In Chapter 3 of this thesis, polynomial functions of the age of different orders were evaluated in the modeling of the random effects associated with the genetic study of the growth curve. In the random part, functions of different orders were considered to model the variances associated with the direct and maternal additive genetic effects and the effects of permanent environment of the animal and the mother. According to the value of the AIC and BIC criterion, the model 6333 allowed to adjust the variance and covariance changes over time and could be used to describe the changes in the variances with the Nelore cattle age. In Chapter 4, we compared different random regression models adjusted by Legendre's orthogonal B-spline and orthogonal segmented polynomials. Although the model that used Legendre's orthogonal polynomials (leg_6333) has been more parsimonious, the model fitting cubic B-spline functions (bspC_6555) with four nodes and three segments for direct additive genetic effects and three nodes and two segments for The maternal genetic effects, the permanent environment of the animal and the permanent maternal environment, was able to adequately describe, with greater speed of convergence, the (co) variance structure of the Nelore bovine weight data set created in the North Brazil.

Keywords: nelore, longitudinal data, random regression models, Legendre, b-splines.

1 INTRODUÇÃO GERAL

O sucesso da bovinocultura de corte depende, em grande parte, da constante melhoria na constituição gênica dos rebanhos. Animais com maior potencial para produção de carne permitem incrementar a produção por meio de técnicas de manejo reprodutivo que visem o melhoramento genético do rebanho. Para isto, deve-se selecionar os melhores animais para serem reprodutores e dessa forma, permitir que os seus genes, responsáveis pelo incremento no potencial genético médio da população, sejam repassados às gerações futuras.

A predição do potencial genético em gado de corte pode ser feita com base na análise da sua curva de crescimento, e essa curva é gerada pela regressão do peso do animal em função da idade. Por meio de medições sucessivas do peso dos animais é possível construir uma trajetória média da curva de crescimento da população, utilizando modelos de regressão aleatória.

Nos modelos de regressão aleatória são utilizados dados do animal e de seus parentes, sendo possível assim, decompor a trajetória da curva por meio das mudanças nas covariâncias dos efeitos aleatórios além de se estimar parâmetros genéticos para todas as idades e etapas da vida do animal.

Diferentes tipos de polinômios podem ser usados para ajustar as funções de covariância nos modelos de regressão aleatória. O tipo e a ordem de ajuste dos polinômios utilizados para modelar as trajetórias médias de crescimento podem influenciar os resultados das avaliações genéticas, portanto, a utilização de modelos estatísticos e metodologias de avaliação genética inadequadas levam à diminuição da acurácia e podem reduzir o progresso genético.

Quando são utilizados polinômios ortogonais para ajuste das funções de covariância, muitas vezes, são necessárias elevadas ordens polinomiais, e conseqüentemente, um grande número de parâmetros a serem estimados. Segundo Meyer (2005), polinômios de alta ordem apresentam maior exigência computacional e instabilidade numérica, colocam grande ênfase nas observações dos extremos da curva (fenômeno “Runge”), levando a estimativas errôneas.

O uso de polinômios segmentados do tipo B (Funções B-spline) tem sido proposto como alternativa aos polinômios de alta ordem, no ajuste das funções de covariância, por permitir unir vários segmentos de polinômios de baixo grau em posições específicas, chamadas “nós”, para formar uma só curva contínua. Em comparação com os polinômios ortogonais, o uso de polinômios B-splines apresentaram vantagens por melhor ajustar os extremos da curva e por sua convergência ser mais rápida (BOLIGON et al., 2012). Segundo Menezes et al. (2010) o uso de polinômios B-splines vem despertando maior interesse por apresentar resultados satisfatórios na modelagem dos efeitos fixos e aleatórios que compõem a equação da curva de crescimento.

Aplicações práticas dos modelos de regressão aleatória ajustados com polinômios segmentados têm ganhado destaque na modelagem de crescimento animal. O maior desafio para a utilização destas funções segmentadas é a identificação adequada do número de intervalos necessários e a posição dos nós na curva, ou seja, o comprimento de cada segmento. Novos estudos sobre aplicação de funções B-spline em modelos de regressão aleatória são necessários para uma melhor avaliação genética dos bovinos de corte.

O objetivo desta pesquisa foi comparar modelos de regressão aleatória ajustados por diferentes funções polinomiais e, com isso, estimar o modelo mais adequado para descrever mudanças nos parâmetros de avaliação do crescimento de bovinos da raça Nelore Mocho criados na região Norte do Brasil.

Esta Tese foi elaborada segundo as normas para elaboração e apresentação de Tese do Programa de Pós-Graduação em Ciência Animal da Universidade Federal do Piauí - (PPGCA/UFPI), dessa forma, os capítulos seguintes serão apresentados no formato de artigo, sendo a formatação de cada artigo feita de acordo com o modelo de publicação definido pelo periódico escolhido.

2 REVISÃO DE LITERATURA

2.1 A bovinocultura de corte no Brasil

O Brasil é o maior exportador e o segundo maior produtor de carne bovina do mundo (USDA, 2016). Portanto, a pecuária de corte é um dos pilares do agronegócio brasileiro e apresenta grande destaque na economia nacional.

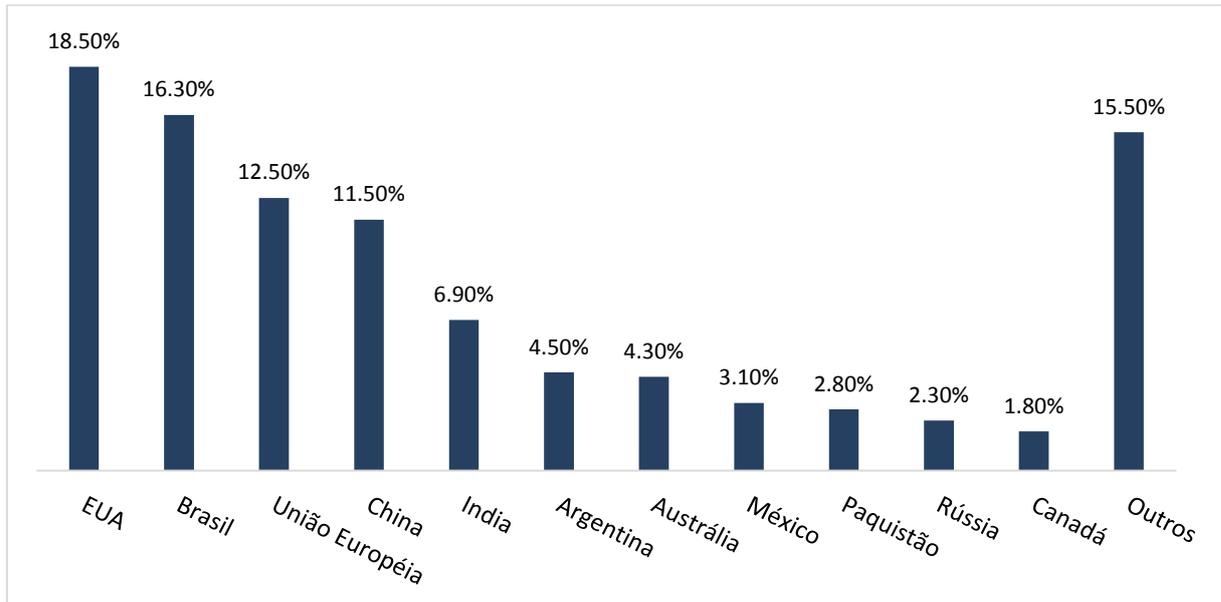


Figura 1. Produção mundial de carne bovina, em porcentagem (USDA, 2016).

Apesar da grande importância econômica para o País, a cadeia produtiva da carne ainda opera com baixa eficiência.

O Brasil possui grande quantidade de terras destinada à pastagem, em torno de 170 milhões de hectares, com a produção de grãos se superando a cada ano (CONAB, 2015), porém, a produtividade relativa nos rebanhos de corte nacional ainda é baixa. Um dos motivos para a baixa produtividade é o uso de animais sem aptidão para produção de carne ou com potencial genético inadequado para a região onde são criados (MAPA, 2014). Essa ineficiência reflete em desperdício de insumos na produção animal.

Da população de bovinos criados no Brasil, cerca de 80% é composta por animais de espécie zebuína (*Bos indicus*), reconhecidos por serem animais rústicos, adaptados ao clima

tropical, e por apresentarem maior resistência a endo e ecto parasitas. Dentre as raças da espécie zebuínas destinadas à pecuária de corte no Brasil, a de maior importância é a raça Nelore, que conta com um efetivo correspondente a 90% desta parcela de zebuínos (ABCZ, 2016).

Apesar de, originalmente, ter sido utilizada para a exploração leiteira, a raça Nelore passou por intenso melhoramento genético, sendo direcionado, quase que exclusivamente, à produção de carne (EUCLIDES FILHO et al., 2004).

A raça Nelore é caracterizada por apresentar ossatura leve, robusta e forte, com musculatura compacta e bem distribuída. O gado Nelore pode apresentar ausência de chifres (Mocho), sendo permitido batoque ou calo, sem protuberância córnea (ACNM, 2016).

Durante muitos anos, a raça Nelore apresentou um número reduzido de animais, porém, nas últimas décadas essa raça se expandiu surgindo variedades, dentre elas a Mocha que assume hoje o segundo lugar em número de registros genealógicos (AMARAL et al., 2014). O gado Nelore Mocho começou a ser registrado como raça pela Associação Brasileira dos Criadores de Zebu em 1961 (ABCZ, 2016). De acordo com alguns pesquisadores, o fator Mocho vem do gado nacional, descendente de bovinos de origem europeia (FARIA et al., 2002).

Com base no expressivo número de animais, a pecuária da região Norte do Brasil é composta, em sua maioria, por animais Nelore, puros ou mestiços, criados a pasto. Como consequência de secas periódicas, baixo potencial genético dos animais e das pastagens, os índices produtivos destes animais são baixos e têm como um dos motivos a falta de programas de melhoramento genético abrangentes (MALHADO et al., 2005).

Nesse contexto, a raça Nelore Mocho merece uma atenção voltada à avaliação e melhoria do desempenho produtivo quanto à precocidade e rendimento de carne. Suas características desejáveis pelo mercado consumidor e, sobretudo, por sua grande adaptação às condições edafoclimáticas da região Norte, torna relevante a adoção de ferramentas tecnológicas capazes de incrementar a produção como a utilização de critérios de melhoramento genético animal,

principalmente o uso de métodos de seleção e cruzamentos, bem como a união destes (BUENO et al., 2011).

Para a seleção dos melhores animais pelo desempenho produtivo em uma característica, deve-se levar em conta os parâmetros genéticos da mesma, além das correlações com as demais características de importância econômica.

As características de crescimento corporal estão relacionadas, principalmente, à precocidade e à qualidade de carcaça, e têm grande destaque em programas de melhoramento por estarem ligadas diretamente ao ganho econômico para o produtor e atenderem às necessidades do mercado consumidor (PAZ et al., 1999). São características que podem ser mensuradas repetidas vezes no decorrer da vida do animal, e por isso são chamadas de medidas repetidas ou longitudinais.

Dentre as características de interesse no progresso genético, o peso é uma medida longitudinal adotada para o monitoramento do crescimento de bovinos de corte, e pode ser obtida em larga escala a um custo baixo (DIAS et al., 2006)

2.2 Modelagem do crescimento animal

Dados de crescimento mensurados em um mesmo indivíduo, do nascimento até a fase adulta, descrevem uma trajetória curva, chamada de curva de crescimento animal. Dentre as medidas corporais utilizadas para a modelagem de crescimento, uma das mais comuns e considerada de fácil mensuração é o peso do animal em determinadas idades. Segundo Albuquerque et al. (2007), as características de peso podem ser aferidas em grandes populações por um custo relativamente baixo.

A regressão do peso em função da idade do animal, por meio de sucessivas pesagens ao longo do tempo, pode ser agrupada em parâmetros que são necessários para a análise genética, que servem de norte para a tomada de decisão na atividade (SANTORO et al., 2005).

O método de análise de dados longitudinais é de grande relevância para o melhoramento genético. Nessas análises, diversos fatores devem ser levados em conta, como: raça (ou grupos genéticos), sexo, ano de nascimento, manejo (entre outros efeitos sistemáticos), como também, o potencial individual, atribuído ao valor genético aditivo de cada indivíduo. Portanto, considerando-se que o fenótipo de um animal é a soma de efeitos fixos e aleatórios, a magnitude desses efeitos, bem como os parâmetros a ele atribuídos, também se modificam à medida que o animal fica mais velho (FREITAS, 2005).

A análise da curva de crescimento permite decompor a variância fenotípica identificando variações decorrentes de fatores fixos e aleatórios, principalmente a variância genética aditiva, variância de ambiente e a variância residual, considerando as diferentes classes de idade, as quais caracterizam as fases do desenvolvimento corporal.

Tendo em vista a padronização das análises genéticas, em programas de melhoramento são estabelecidas idades nos quais os pesos são ajustados. Essas idades são definidas como idades-padrão, que contemplam princípios biológicos como por exemplo: peso ao nascimento, desmama, ano, sobreano e abate (SAKAGUTI et al., 2003). Entretanto, essas pesagens são difíceis de serem padronizadas durante a coleta, dificultando o processo de avaliação genética.

Para análise de regressão do peso do animal sobre sua idade, através de medidas repetidas no tempo, são indicados os modelos de regressão aleatória por propiciar a curva média de crescimento dos animais (SOUSA JUNIOR et al., 2010). Uma das vantagens do uso de modelos de regressão aleatória é que permitem identificar as variações decorrentes de fatores fixos e aleatórios em qualquer idade do animal, sendo dispensável o uso de idades-padrão para análises genéticas.

2.3 Modelos de regressão aleatória e funções de covariância

Os modelos de regressão aleatória e as funções de covariância podem expressar, de maneira mais realista, os fenômenos associados a dados longitudinais, por exemplo, curvas de crescimento através do tempo ou medições repetidas (RESENDE et al., 2001).

Com a evolução das ferramentas computacionais, os modelos de regressão aleatória têm se tornado a metodologia padrão para análises de medidas longitudinais no melhoramento animal (PALHARIM et al., 2013). Entre outras vantagens, a utilização desses modelos permite o uso de forma eficiente de toda a informação disponível de cada animal, já que todas as medidas do animal e de seus parentes são utilizadas para a avaliação genética (MEYER, 2005), e, com isso, aumenta a acurácia das estimativas dos valores genéticos preditos.

Diferente dos modelos de repetibilidade e multivariado, nos modelos de regressão aleatória não há a necessidade de se utilizar fatores de ajuste para idades-padrão, uma vez que estes já estão incluídos por meio da matriz de delineamento. Neste contexto, os modelos de regressão aleatória adequados para explicar melhor a variabilidade associada a medidas repetidas do que os modelos lineares simples, pois permitem a obtenção de diferentes curvas de valores genéticos associadas aos diferentes indivíduos e considera as mudanças nas variâncias genéticas e residuais através do tempo. Também, desde que existam estimativas confiáveis de componentes de variância, estes modelos permitem a predição de valores genéticos de um indivíduo em diferentes idades, baseando-se em avaliações de apenas uma idade (RESENDE et al., 2001). Isso é possível porque os modelos de regressão aleatória consideram as mudanças na média e variâncias da característica ao longo do tempo, geradas pelas funções de covariâncias.

As funções fornecem as (co)variâncias de características medidas em diferentes pontos de uma trajetória, descrevendo a (co)variância entre medidas tomadas em determinadas idades como uma função destas idades (VAN DER WERF & SCHAEFFER, 1997). Estas funções, quando estimadas por regressão aleatória, são úteis para análise de dados de crescimento dos

animais Nelore e aplicadas visando o progresso genético destes (BOLIGON et al., 2010; SANTOS et al., 2011; PALHARIM et al., 2013).

Ao se modelar a curva de variância como uma regressão em função de uma covariável contínua, os modelos de regressão aleatória são facilmente implementados em análises de modelos lineares mistos, considerando a influência de efeitos fixos e aleatórios (MEYER, 2005). Em avaliações genéticas da curva de crescimento, o ajuste de uma função contínua fixa é denominada de trajetória média de crescimento (KIRKPATRICK, M.; LOFSVOLD, D.; BULMER, M., 1990). Ela representa a tendência de crescimento do peso corporal médio da população. As curvas relativas ao efeito aleatório, são, portanto, variações individuais dos animais que indicam o seu efeito genético e, de um modo geral, essas variações podem ser entendidas como desvios da produção dos animais em torno da curva fixa (SILVA et al., 2011).

Os modelos de regressão aleatória consideram diferentes estruturas de heterogeneidade de variância residual por meio da formação de classes de resíduo ao longo das idades de avaliação coletadas. A importância de se considerar a variância residual heterogênea reside no fato de que se pode melhorar a partição da variância total, o que proporciona a estimação de parâmetros e valores genéticos mais acurados. Segundo El Faro e Albuquerque (2005), esse tipo de modelagem da variância residual pode aumentar o número de parâmetros a serem estimados no processo de maximização da função de verossimilhança.

O número elevado de parâmetros no modelo, a estrutura de (co)variância e o fato de as matrizes de incidência serem mais densas que as dos modelos convencionais podem dificultar a convergência na estimação dos componentes de variância, sendo entraves para a utilização dessa metodologia para avaliações genéticas em situações com grande número de dados.

As funções de covariância são modeladas por polinômios, e de modo geral, do tipo ortogonal. No entanto, polinômios alternativos como os segmentados, ou também conhecidos

por funções Splines, podem ser utilizados em modelos de regressão aleatória apresentando um melhor ajuste à curva (BALDI, F.; ALENCAR, M. M.; ALBUQUERQUE, L. G., 2010).

2.4 Funções B-splines

Funções Splines ou polinômios segmentados são curvas formadas por inúmeros segmentos de polinômios de menor grau, que se unem entre si por meio de posições determinadas, denominadas nós, formando assim uma curva contínua (BALDI, F.; ALENCAR, M. M.; ALBUQUERQUE, L. G., 2010). Tais nós (pontos de união) apontam o local em que ocorre a mudança de comportamento da curva, esse fenômeno é denominado de grau de polinômio ou inclinação (PAZ et al., 1999).

Existe ainda uma extensão das funções Splines, conhecida por funções B-splines, que assim como a primeira, também é formada por segmentos ou “pedaços” de polinômios, sendo que estas compõem uma base de espaços Splines, apresentando como duas principais vantagens, a alta flexibilidade e rápida computação (FROZZA, 2010). Funções do tipo B-Splines possuem melhores propriedades numéricas, além de serem indicadas na modelagem de efeitos aleatórios, quando analisados com uso de modelos mistos, bem como para estimação de funções de covariância (BRITO, 2012),

O maior desafio na utilização destas funções é identificar a localização e a quantidade de nós para uma modelagem satisfatória da curva. Segundo Scalez et al. (2011) o polinômio B-spline quadrático com quatro intervalos se ajustou adequadamente à trajetória média de crescimento de tourinhos Nelore, porém, quando avaliadas as trajetórias médias de crescimento de tourinhos mestiços Nelore e Canchim, o melhor ajuste foi obtido com uso de polinômios quadráticos do tipo B-splines com seis intervalos.

Segundo Menezes (2010), as funções B-Splines podem ser utilizadas de diversos modos, por meio de diferentes bases, sendo que em virtude do modo escolhido, essas funções podem

reduzir a multicolinearidade, além de apresentarem rápida convergência quando comparado aos modelos mais parametrizados.

Tholon et al. (2012), descreveram um modelo geral de um polinômio segmentado:

$$Y = a_0 + a_1X + a_2x^2 + b_1Z_1 + b_2Z_2 + b_3Z_3 + b_4Z_4$$

Em que:

$$Z_1 = (X - K_1)^2, \text{ se } X > K_1;$$

$$Z_2 = (X - K_2)^2, \text{ se } X > K_2;$$

$$Z_3 = (X - K_3)^2, \text{ se } X > K_3;$$

$$Z_4 = (X - K_4)^2, \text{ se } X > K_4.$$

Y = é o peso, X = é a idade, a_0 = é o intercepto, a_1 e a_2 = são os parâmetros linear e quadrático, b_1 , b_2 , b_3 e b_4 = são os coeficientes de regressão dos polinômios e K_1 , K_2 , K_3 e K_4 = representam os nós.

O emprego de polinômios segmentados B-splines nos estudos genéticos do crescimento de bovinos busca esclarecer a ordem apropriada de ajuste das funções polinomiais para os diferentes efeitos aleatórios e a melhor estrutura de variância para modelar os resíduos.

3 REFERÊNCIAS BIBLIOGRÁFICAS DA INTRODUÇÃO GERAL E REVISÃO

- ASSOCIAÇÃO DOS CRIADORES DE NELORE MOCHO – ACNM. Disponível em: <www.neloremocha.com.br/index.php?id=Raça> Acesso em 30/11/2016.
- ASSOCIAÇÃO BRASILEIRA DOS CRIADORES DE ZEBU – ABCZ. Estatísticas. [2016] Disponível em: < <http://www.abcz.org.br/Home/Conteudo/23985-Racas-Zebuinas> > Acesso em 30/11/2016
- ALBUQUERQUE, L. G; MERCADANTE, M. E. Z; ELER, J. P. Aspectos da seleção de *Bos indicus* para produção de carne. **Boletim de Indústria animal**, v.64, p.339-348, 2007.
- AMARAL, R. D. S., CARNEIRO, P. L. S., AMBROSINI, D. P., & MALHADO, C. H. M. Parâmetros Populacionais do Rebanho Nelore Mocho Registrado no Nordeste Brasileiro. **Revista Científica de Produção Animal**, v. 13, n. 1, p. 54-57, 2011.
- BALDI, F.; ALENCAR, M. M.; ALBUQUERQUE, L. G. Random regression analyses using B-splines functions to model growth from birth to adult age in Canchim cattle. **Journal of Animal Breeding and Genetics**, v.127, n.6, p.433-441, 2010.
- BOLIGON, A. A.; SILVA, J. A. V.; SESANA, R. C.; SESANA, J. C.; JUNQUEIRA, J. B.; ALBUQUERQUE, L. G. Estimation of genetic parameters for body weights, scrotal circumference, and testicular volume measured at different ages in Nelore cattle. *Journal of Animal Science*, Champaign, v. 88, n. 4, p. 1215-1219, 2010
- BOLIGON, A. A. et al. Random regression analyses using B-spline functions to model growth of Nelore cattle. **Animal**, v.6, n.2, p.212-220, 2012
- BRASIL. Ministério da Agricultura, Pecuária e Abastecimento. **Plano mais pecuária**. Brasília: MAPA/ACS, 2014. 32p.
- BRITO, L. F. **Modelagem da produção de leite de cabras das raças Alpina e Saanen utilizando modelos de regressão aleatória**. 2012. 82f. Dissertação (Mestrado em Genética e Melhoramento) – Universidade Federal de Viçosa, Viçosa, MG.
- BUENO, R. S. et al. Inclusão da epistasia em modelo de avaliação genética de bovinos de corte compostos. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.63, p.948-953, 2011.
- CONAB. Indicadores da Agropecuária: **Quadro de Suprimentos**. Disponível em <<http://www.conab.gov.br/conteudos.php?a=1470&t=2>> Acesso em junho de 2016.
- DIAS, L. T. et al. Estimação de parâmetros genéticos para peso do nascimento aos 550 dias de idade para animais da raça Tabapuã utilizando-se modelos de regressão aleatória. **Revista Brasileira de Zootecnia**, v.35, n.5, p.1915-1925, 2006.
- EILERS, P. H. C.; MARX, B. D. Flexible smoothing with B-splines and penalties. **Statistical Science**, v. 11, p.89-121, 1996.
- EL FARO, L.; ALBUQUERQUE, L. G. Predição de valores genéticos para a produção de leite no dia do controle e para produção acumulada até 305 dias. **Revista Brasileira de Zootecnia**, v.34, n.2, p.496-507, 2005.

EUCLIDES FILHO, K; et al. Eficiência bionutricional de animais da raça Nelore, F1s Valdostana-Nelore e de mestiços de raças europeias adaptadas. **Arquivo Brasileiro Medicina Veterinária e Zootecnia**, v.56, n.5, p.671-675, 2004

FARIA, F.J.C.; VERCESI-FILHO, A.E.; MADALENA, F.E.; JOSAHKIAN, L.A. Estrutura populacional da raça Nelore Mocho. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.54, p.501-509, 2002.

FREITAS, A. R. Curvas de crescimento na produção animal. **Revista Brasileira de Zootecnia**, v.34, n.3, p.786-795, 2005.

FROZZA, M. **Introdução à Análise Descritiva de Dados Funcionais**. 2010. 57f. Monografia (Bacharelado em Estatística) - Universidade Federal do Rio Grande do Sul, Porto-Alegre.

KIRKPATRICK, M.; LOFSVOLD, D.; BULMER, M. Analysis of the inheritance, selection and evolution of growth trajectories. **Genetics**, v.124, n.4, p.979-993,1990.

MALHADO, C.H.M.; MARTINS-FILHO, R.; LÔBO, R.N.B.; et al. Tendências genéticas para características relacionadas à velocidade de crescimento em bovinos Nelore na região Nordeste do Brasil. **Revista Brasileira de Zootecnia**, v.34, p.60-65, 2005

MENEZES, G. R. O. **Uso de polinômios segmentados na modelagem de dados longitudinais de ponderal em bovinos da raça Tabapuã**. 2010. 99f. Tese (Doutorado em Genética e Melhoramento) - Universidade Federal de Viçosa, Viçosa, MG.

MEYER, K. Random regression analyses using B-splines to model growth of Australian Angus cattle. **Genetics Selection Evolution**, v.37,n.5, p.473-500, 2005.

MEYER, K. WOMBAT— Digging deep for quantitative genetic analyses by restricted maximum likelihood. **Proceedings...** World Congress on Genetic and Applied Livestock Production, Belo Horizonte, Brasil, p.27-14, 2006.

PALHARIM, D. A.; ARAÚJO, C. V.; BITTENCOUR, T. C. B. S. C.; ARAÚJO, S. I.; LÔBO, R. B.; BEZERRA, L. A. F. Estimativas de herdabilidade para o desempenho ponderal na raça Nelore no Mato Grosso. **Scientific Electronic Archives**, Sinop, v. 3, n. 1, p. 59-62, 2013.

PAZ, C. C. P.; ALBUQUERQUE, L. G.; FRIES, L. A. Efeitos ambientais sobre ganho de peso no período do nascimento ao desmame em bovinos da raça Nelore. **Revista Brasileira de Zootecnia**, v.28, n.1, p.55-64, 1999.

RESENDE, M. D. V.; REZENDE, G. D. S. P.; FERNANDES, J. S. C. Regressão aleatória e funções de covariância na análise de medidas repetidas. **Revista de Matemática e Estatística**, v.19, s/n, p.21-40, 2001.

SAKAGUTI, E. S. et al. Avaliação do crescimento de bovinos jovens da raça Tabapuã, por meio de análises de funções de covariâncias. **Revista Brasileira de Zootecnia**, v.32, n.4, p.864-874, 2003.

SANTORO, K.R.; BARBOSA, S.B.P.; SANTOS, E.S.; BRASIL, L.H.A. Herdabilidades de parâmetros de curvas de crescimento não-lineares em zebuínos, no Estado de Pernambuco. **Revista Brasileira de Zootecnia**, v.34, p. 2280-2289, 2005.

SANTOS, G. C. J.; LIRA, T. S.; PEREIRA L. S.; LOPES, F. B.; FERREIRA, J. L. Efeitos não genéticos sobre características produtivas em rebanhos Nelore criados na região norte do Brasil. **Acta Veterinária Brasilica**, Mossoró, v. 5, n. 4, p. 385-392, 2011.

SCALEZ, D. C. B. et al. Polinômios parTabela 0.1a modelar a trajetória de crescimento de tourinhos em provas de ganho em peso. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ZOOTECNIA, 48, 2011, Belém. **Anais...** Belém: Sociedade Brasileira de Zootecnia, 2011.

SILVA, F. L. et al. Curvas de crescimento em vacas de corte de diferentes tipos biológicos. **Pesquisa Agropecuária Brasileira**, Brasília, v. 46, n. 3, p. 262-271, 2011.

SOUSA JÚNIOR, S. C. et al. Estimação de funções de covariância para características de crescimento da raça Tabapuã utilizando modelos de regressão aleatória. **Revista Brasileira de Zootecnia**, v.39, n.5, p.1037-1045, 2010.

THOLON, P. et al. Utilização de funções lineares e não lineares para ajuste do crescimento de bovinos Santa Gertrudis, criados a pasto. **Ars Veterinaria**, v.28, n.1, p.234-239, 2012.

USDA. **USDA Foreign Agricultural Service**. Disponível em <<http://www.fas.usda.gov/regions/brazil>>. Acesso em Novembro de 2016

VAN der WERF, J.; SCHAEFFER, L. **Random Regression in Animal Breeding**. Course Notes, Ontario: University of Guelph, 1997, 70p.

4 CAPÍTULO I

Modelagem da curva fixa de crescimento de bovinos Nelore Mocho da região Norte do Brasil

Resumo: Funções polinomiais da idade com diferentes ordens foram avaliadas na modelagem da trajetória média de crescimento em modelos de regressão aleatória. Utilizaram-se 15.148 registros de pesos de 3.115 bovinos da raça Nelore Mocho com idade entre 1 e 660 dias, criados na região norte do Brasil e nascidos entre os anos de 1995 e 2010. Inicialmente, a trajetória média da população foi ajustada por uma regressão fixa sob polinômios ortogonais da idade com ordens variando de dois a sete. Estas análises foram executadas por meio do método de quadrados mínimos ordinários, desconsiderando os efeitos aleatórios do animal e materno. Posteriormente, para observar o impacto de diferentes ordens de ajustes da regressão fixa na estimativa dos parâmetros genéticos, as regressões fixas e aleatórias foram ajustadas por meio de polinômios de Legendre. Os efeitos fixos considerados foram os de grupos de contemporâneos e a idade da mãe ao parto, como covariável linear e quadrática. Na parte aleatória do modelo incluiu-se os efeitos aleatórios de animal e materno (genéticos diretos e ambientes permanentes). O aumento na ordem do polinômio do modelo melhorou o ajuste aos dados observados, segundo os critérios de quadrado médio do resíduo (QMR), desvio médio absoluto (DMA), coeficiente de determinação (R^2) e percentual quadrado de viés (PQV). Para ambas as idades, 240 e 660 dias, a correlação de Spearman foi máxima (100%) entre os valores genéticos preditos a partir da função cúbica (ordem 4). Uma função polinomial de ordem quatro (cúbica) pode ser utilizada para descrever a curva média de crescimento de bovinos da raça Nelore Mocho sem prejuízo para a análise dos valores genéticos em ambas idades.

Palavras-chave: curva fixa, dados longitudinais, modelos de regressão aleatória, Legendre, Nelore.

CHAPTER I

Modeling of the Nelore Mocho bovine growth curve in the Northern region of Brazil

Abstract: polynomial functions of age with different orders were evaluated in modeling the average growth trajectory in random regression models. We used 15,148 weight records of 3115 Nelore Mocho cattle aged between 1 and 660 days, raised in northern Brazil born between 1995 and 2010. Initially, the average trajectory of the population was adjusted by a fixed regression on orthogonal polynomials of age with orders ranging from two to seven. These analysis were performed using the method of ordinary least squares, disregarding the animal and maternal random effects. Subsequently, to observe the impact of different orders fixed regression fits in the genetic parameters, fixed and random regressions were fitted by Legendre polynomials. The increase in the order of the polynomial of the model improved fit to observed data, according to the criteria of mean square of error (MSE), average absolute deviation (AVG), coefficient of determination (R^2) and percentage quadratic of bias (PQV). For both age 240 to 660 days, Spearman correlation was maximum (100%) between the predicted values from the function cubic (order 4). A polynomial function of order equal to four (cubic) can be used to describe the average growth curve of Nelore Mocho breed cattle without prejudice to the analysis of genetic values at both ages.

Keywords: fixed curve, longitudinal data, random regression models, Legendre, Nelore

Introdução

Os modelos de regressão aleatória permitem o ajuste de curvas aleatórias para cada indivíduo incluído no modelo, as quais são expressas como desvios de uma curva média da população ou grupo contemporâneo ao qual pertence o indivíduo. Ao se ajustar modelos de regressão aleatória assume-se que há certa estrutura de covariâncias entre as observações, e esta é determinada pelas covariâncias entre os coeficientes da regressão, podendo essa equação ser caracterizada como uma função de covariâncias (Meyer, 1998).

Nos estudos das curvas de crescimento animal, a função contínua da parte fixa dos modelos de regressão aleatória é conhecida como curva média de crescimento, por representar a tendência de crescimento do peso corporal médio da população em relação às diferentes idades (Kirkpatrick et al., 1990). Portanto, deve ser dada a devida atenção à modelagem da parte fixa do modelo, pois esta tem influência na estimação dos parâmetros genéticos e, conseqüentemente, na conformidade dos objetivos de seleção.

Para modelagem adequada da curva em bovinos de corte, alguns autores investigaram não somente a ordem de ajuste de funções polinomiais para representar a parte aleatória, como também a curva média de crescimento em modelos de regressão aleatória. Meyer (1999) estudou o crescimento pós-desmame de fêmeas de corte de duas raças bovinas na Austrália e comparou por meio de polinômios ortogonais da idade, funções com ordens variando de dois até oito. Nesse estudo a autora concluiu que o crescimento dos animais foi melhor descrito por uma regressão de ordem quatro (cúbica). Posteriormente, Meyer (2003) sugeriu que uma regressão quadrática para todos os efeitos aleatórios seria uma escolha conservadora e sensível.

No Brasil, Santoro et al. (2005), estudando o uso de funções de covariância na descrição do crescimento de bovinos Nelore criados no Estado do Pernambuco, concluíram que o modelo que considerou tanto os efeitos fixos quanto os aleatórios com polinômios de ordem 5 foi o mais adequado ao ajuste dos dados. Sakaguti et al. (2002) estudando o crescimento de bovinos Tabapuã, observaram que a curva média de crescimento deve ser representada por polinômios de, no mínimo, quarta ordem (cúbica). Seguindo a mesma metodologia, Baldi et al. (2010) estudando o crescimento de bovinos de corte da raça Canchim utilizaram polinômios de Legendre de ordem quatro para modelagem da curva média e diferentes ordens para ajuste dos efeitos aleatórios.

O incremento na ordem de ajuste para a regressão da curva fixa deve ser feito com cautela, observando-se a significância das mesmas para o ajuste da curva fixa, pois a utilização de

funções com menores ordens é preferível. Nesse aspecto, Arango et al. (2004) afirmam que a regressão quadrática da idade (ordem 3) foi a ordem polinomial mais alta com efeito significativo, e isto foi determinante para os autores assumirem como a ordem de melhor ajuste. No entanto, ao contrário dos trabalhos citados anteriormente, Cobucci et al. (2011) não encontrou diferenças significativas nos parâmetros genéticos estimados ao comparar diferentes ordens de ajustes na modelagem da curva fixa em modelos de regressão aleatória.

Diante das divergências encontradas na literatura a respeito da modelagem da curva fixa em modelos de regressão aleatória, se faz necessário realizar estudos sobre a curva fixa na modelagem de crescimento de bovinos de corte da raça Nelore Mocho.

Com este trabalho, objetivou-se avaliar a utilização de funções polinomiais na descrição da curva média de crescimento de bovinos da raça Nelore Mocho com finalidade de estimar a melhor ordem para ajuste dos efeitos fixos em avaliações genéticas utilizando modelos de regressão aleatória, bem como verificar a influência de diferentes ajustes da parte fixa dos modelos nas estimativas de parâmetros genéticos.

Material e Métodos

Descrição dos dados

Os dados são provenientes de rebanhos participantes do Programa de Desenvolvimento Ponderal da Associação Brasileira de Criadores de Zebu (ABCZ). Estes rebanhos estão localizados em sete estados da região Norte do Brasil (Amazonas, Acre, Amapá, Roraima, Rondônia, Pará e Tocantins).

Inicialmente, no banco de dados constavam 44.106 registros. A edição dos dados consistiu em eliminar: registros de pesos com idade acima de 660 dias; animais sem identificação do pai e da mãe; animais com menos de 3 pesagens; e grupos contemporâneos com menos de 3 animais.

A definição de grupo contemporâneo incluiu: sexo, ano de nascimento, estação do ano ao nascimento, ano da pesagem, mês da pesagem, fazenda e condição de criação. As estações de nascimento foram classificadas em chuvosa (dezembro a maio) e seca (junho a novembro). As condições de criação foram: “mamando” e “desmamado”, com regime alimentar à pasto.

Uma análise de consistência dos dados foi feita com finalidade de excluir registros com possíveis erros de medição (*outliers*), através da constatação de uma variabilidade expressiva

dentro de cada grupo contemporâneo. Foram excluídos os registros de peso que apresentavam valores fora do intervalo de ± 3 desvios-padrão em relação à média do grupo.

Após edição dos dados, havia 15.148 observações de peso pertencentes a 3.115 animais nascidos entre os anos de 1995 a 2010 e criados em regime alimentar à pasto, com peso médio ao nascer de 28 kg. O peso médio e o número de registros por idade estão representados na figura 1.1.

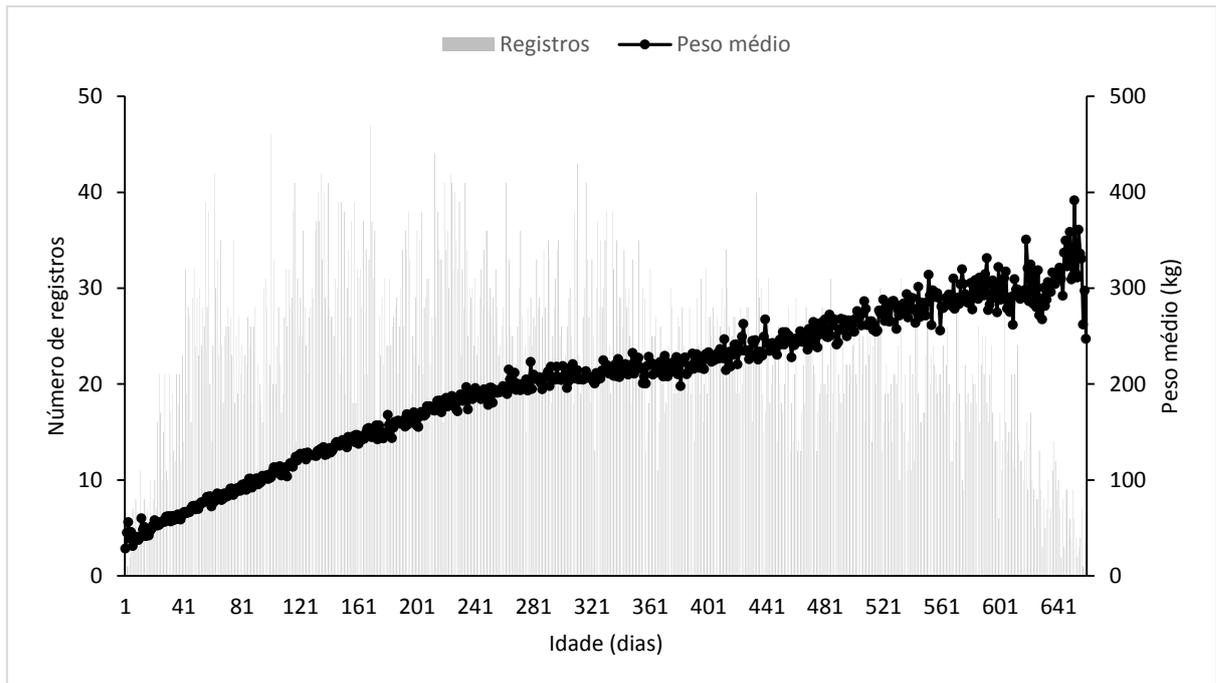


Figura 1. 1. Número de registro e peso corporal médio por idade de bovinos Nelores provenientes da região Norte do Brasil.

A estrutura de dados com número de observações por animal e informações referentes ao pedigree está representada na Tabela 1.1.

Tabela 1. 1. Resumo da estrutura dos dados de bovinos Nelores provenientes da região Norte do Brasil utilizados no estudo.

Informação	Quantidade
Dados	15.148
Animais com observação	3.115
Com 3 observações	712
Com 4 observações	698
Com 5 observações	549
Com 6 observações	659
Com 7 a 10 observações	497
Reprodutores	282
Matrizes	764

Animais na matriz de parentesco	5.678
Grupos contemporâneos	1.819

A edição dos dados e as análises de consistência foram feitas com auxílio do software SAS, versão 9.0 (2003).

Modelos e análises

Inicialmente, a trajetória média da população foi ajustada por uma regressão fixa sob polinômios ortogonais da idade com ordens variando de dois a sete. As análises foram executadas por meio do método de quadrados mínimos ordinários, desconsiderando os efeitos aleatórios do animal e da mãe. Os demais efeitos fixos considerados nesta análise foram os de grupos de contemporâneos e a idade da mãe ao parto, como covariável linear e quadrática. Para avaliar a adequabilidade de ajuste das funções de diferentes ordens, empregaram-se os seguintes critérios: quadrado médio do resíduo (QMR), calculado como segue:

$$QMR = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 ,$$

em que, n é o número de observações, Y_i e \hat{Y}_i são os valores observados e estimados, respectivamente; o desvio médio absoluto (DMA), que é dado por:

$$DMA = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i| ,$$

utilizado por Sarmiento et al. (2006), obtido como o somatório dos desvios médios em valores absolutos; a porcentagem quadrada de viés (PQV), proposto por Ali e Schaeffer (1987), determinado por:

$$PQV = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i)^2} \times 100 ,$$

e o coeficiente de determinação (R^2), obtido por meio do quadrado da correlação entre os pesos observados e estimados por uma função, equivalente a:

$$R^2 = 1 - \left(\frac{SQR}{SQTc} \right) ,$$

em que, SQR é a soma de quadrados do resíduo e $SQTc$ é a soma de quadrados total corrigida pela média.

Em uma segunda etapa das análises, para avaliar a influência da ordem de ajuste empregada para representar a trajetória média da população sobre os componentes de variância estimados e os valores genéticos preditos, ajustaram-se diferentes modelos de regressão aleatória, variando-se a ordem de ajuste da regressão fixa de dois a sete, e mantendo a parte aleatória da regressão com a ordem de ajuste quadrática. Tanto as regressões fixas como aleatórias foram representadas por funções contínuas, cujas idades foram descritas em termos de polinômios ortogonais de Legendre, por meio do programa Wombat (Meyer, 2007) versão 2015. Os modelos estudados podem ser representados como segue:

$$Y_{ij} = F + \sum_{m=1}^{k\beta-1} \beta_m \Phi_m + \sum_{m=1}^3 \alpha_{im} \Phi_m + \sum_{m=1}^3 \rho_{im} \Phi_m + \sum_{m=1}^3 \gamma_{im} \Phi_m + \sum_{m=1}^3 \delta_{im} \Phi_m + \varepsilon_{ij}$$

Em que Y_{ij} é o peso no dia j do animal i ; F refere-se a um conjunto de efeitos fixos, incluídos no modelo (grupo de contemporâneos e covariável da idade da vaca ao parto, efeitos linear e quadrático); β_m são os coeficientes de regressão fixos do peso sobre o polinômio de Legendre, com $k\beta$ variando de dois a sete ($k\beta = 2, 3, 4, \dots, 7$) para modelar a trajetória média da população.

Na parte aleatória do modelo temos “ $\alpha_{im}, \rho_{im}, \gamma_{im}, \delta_{im}$,” representando os coeficientes de regressão genético aditivo direto, ambiente permanente do animal, genético aditivo materno e de ambiente permanente materno, respectivamente. A ordem dos polinômios para modelagem dos quatro efeitos aleatórios foi igual a três (quadrática); ε_{ij} é o efeito aleatório residual.

Em notação matricial, o modelo acima e suas respectivas pressuposições podem ser descritos como:

$$\mathbf{Y} = \mathbf{Xb} + \mathbf{Z}_1\mathbf{a} + \mathbf{Z}_2\mathbf{c} + \mathbf{Z}_3\mathbf{m} + \mathbf{Z}_4\mathbf{q} + \mathbf{e}$$

$$E \begin{bmatrix} Y \\ a \\ c \\ m \\ q \\ e \end{bmatrix} = \begin{bmatrix} Xb \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, e \quad \begin{aligned} \text{Var}(a) &= \mathbf{K}_a \otimes \mathbf{A}_{Nd}, \\ \text{Var}(c) &= \mathbf{K}_c \otimes \mathbf{I}, \\ \text{Var}(m) &= \mathbf{K}_m \otimes \mathbf{A}_{Nm}, \\ \text{Var}(q) &= \mathbf{K}_q \otimes \mathbf{I}, \\ \text{Var}(e) &= \mathbf{R}; \end{aligned}$$

Em que, Y é um vetor de N observações referentes a N_d animais; b é um vetor que contém os efeitos fixos e os coeficientes b_m da regressão fixa; a é um vetor $K_a \times N_D$ de coeficientes de regressão aleatória genéticos aditivos diretos, em que $N_D > N_d$ denota o número total de animais na análise, ou seja, na matriz de parentesco (5.678); c é um vetor $K_c \times N_d$ de coeficientes de regressão aleatória ambiente permanente de animal; m é um vetor $K_m \times N_D$ de coeficientes de regressão aleatória genéticos aditivos maternos; q é um vetor $K_q \times N_m$ de coeficientes de regressão aleatória de ambiente permanente materno, sendo N_m igual ao número de fêmeas que têm progênie com dados; e é um vetor de erros aleatórios; X, Z_1, Z_2, Z_3 e Z_4 referem-se às matrizes de incidência dos efeitos fixos, genético aditivo direto, ambiente permanente de animal, genético aditivo materno e ambiente permanente materno, respectivamente. K_a, K_c, K_m e K_q são as matrizes de (co)variâncias entre os coeficientes de regressão aleatórios genético aditivo direto, ambiente permanente de animal, genético aditivo materno e ambiente permanente materno, respectivamente; A é a matriz dos numeradores do coeficiente de parentesco entre os indivíduos; I_{N_d} é uma matriz identidade de dimensão N_d ; I_{N_m} é uma matriz identidade de dimensão N_m ; \otimes é o produto de Kronecker; R é uma matriz diagonal de variâncias residuais.

A variância residual foi considerada homogênea durante toda curva de crescimento. A covariância entre os efeitos genéticos aditivos direto e materno foi assumida como sendo igual a zero. As (co)variâncias entre os coeficientes de regressão aleatória genéticos aditivos direto e materno e ambiente permanente de animal e materno, conforme o modelo ajustado foram estimadas pelo método da máxima verossimilhança restrita (REML), usando software Wombat (Meyer, 1998).

Avaliaram-se as mudanças ocorridas nos componentes de variância e parâmetros genéticos estimados ao longo da curva de crescimento, em função da modelagem da regressão fixa.

Resultados e Discussão

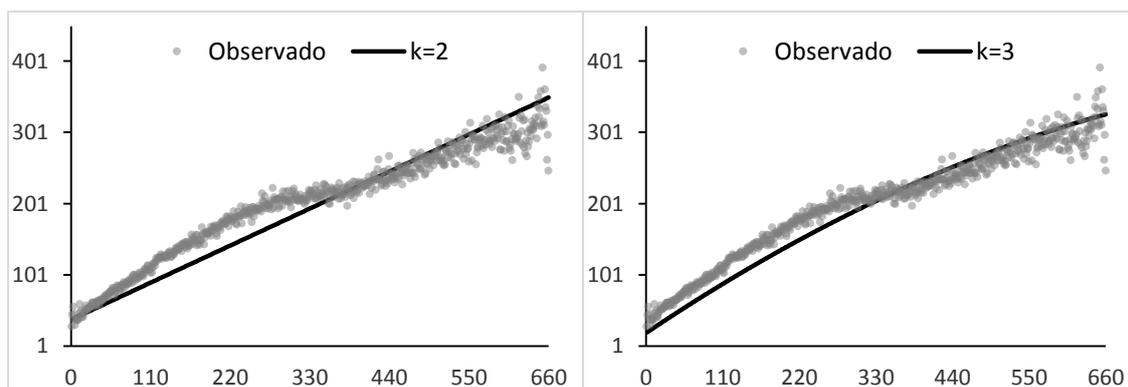
Comparando-se a ordem das funções ajustadas pelo método dos quadrados mínimos ordinários, observou-se que o pior ajuste foi para ordem linear e a maior mudança ocorreu ao passar de um ajuste linear ($k=2$) para quadrático ($k=3$), como indicado pelos critérios utilizados para avaliar qualidade de ajuste, apresentados na Tabela 1.2. Observa-se que, à medida que aumenta a ordem de ajuste, há uma melhoria nos valores de QMR, DMA, R^2 e PQV, sendo que, após a ordem cúbica, as diferenças observadas foram baixas.

Tabela 1. 2. Quadrado médio do erro de predição (QMR), desvio médio absoluto (DMA), coeficiente de determinação (R^2), percentagem quadrada de viés (PQV) para o ajuste da trajetória média da população por uma função polinomial com diferentes ordens (k)

Ordem	QMR ¹	DMA ¹	R^2	PQV
K=2	143,94	8,88	0,97	0,33
K=3	138,31 (-5,62)	8,65 (-0,23)	0,98	0,31
K=4	134,21 (-4,10)	8,53 (-0,12)	0,98	0,31
K=5	133,83 (-0,39)	8,51 (-0,01)	0,98	0,30
K=6	132,46 (-1,37)	8,46 (-0,05)	0,98	0,30
K=7	132,43 (-0,03)	8,46 (0,00)	0,98	0,30

¹ Diferença entre o modelo analisado e o modelo seguinte para os valores de QMR e DMA

Como se observa na Figura 1.2, o ajuste linear foi o pior, apresentando-se distante dos dados no início e final da curva. Nas demais ordens de ajustes, as funções diferiram, principalmente, da metade para o final da curva de crescimento.



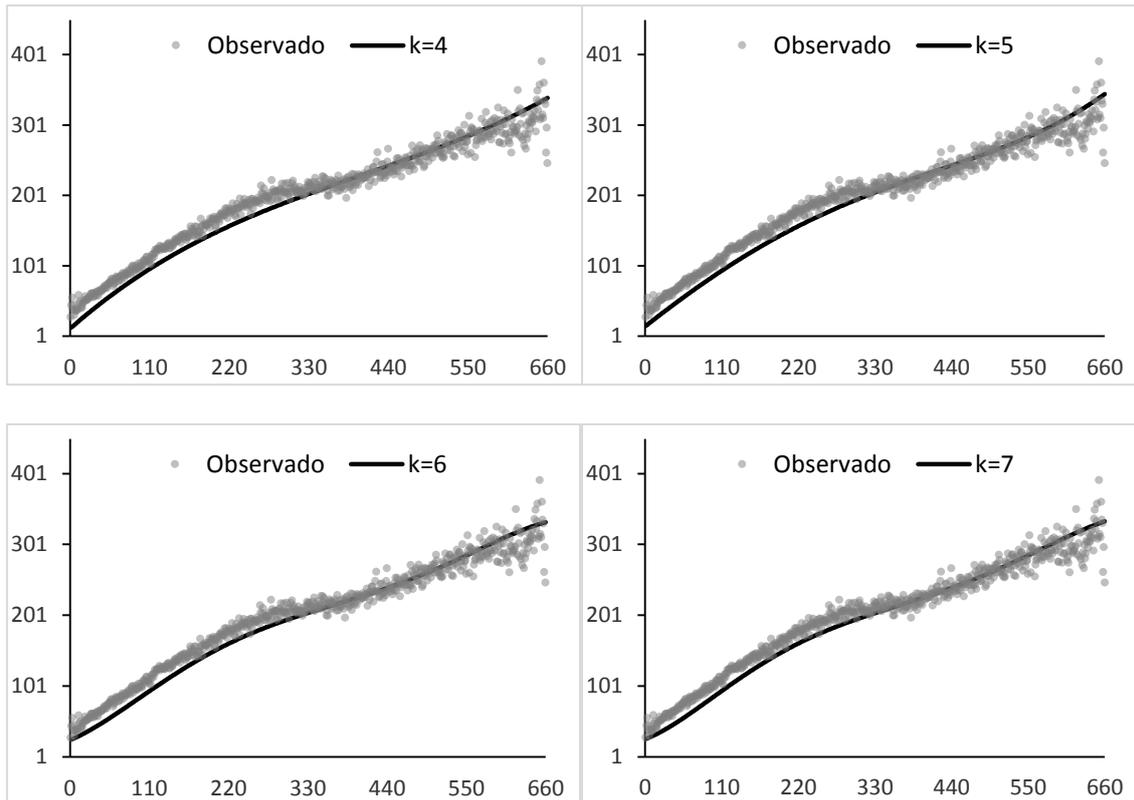


Figura 1. 2. Dispersão dos pesos médios observados em função da idade e trajetória média populacional estimada (regressão fixa sob polinômios ortogonais da idade) por quadrados mínimos de acordo com a ordem de ajuste (K).

As funções com ordem a partir de quatro tenderam a estimar curvas mais próximas entre si, diferenciando-se, principalmente, nos extremos da curva como pode ser observado resultados apresentados na Tabela 1.2 e na Figura 1.2.

Todavia, além de ajustar a curva média de crescimento de forma condizente com o desenvolvimento corporal dos bovinos estudados, as funções com ordem maior que quatro apresentaram maior flexibilidade do meio para o final da curva, nos pontos mais divergentes. Segundo Sakaguti et al. (2002), a menor flexibilidade da função cúbica, em relação às de ordens mais avançadas, parece não ter prejudicado o ajustamento dos modelos de crescimento ao peso corporal dos bovinos, pois segundo os autores, essa menor flexibilidade proporcionou uma curva de crescimento mais condizente com o desenvolvimento corporal dos animais estudados.

A figura 1.3 indica a estimativa de herdabilidade animal, bem como o componente de variância para ambiente permanente animal como proporção da variância fenotípica,

sob diferentes ordens de ajuste da curva média. A variância residual nesses modelos foi considerada homogênea.

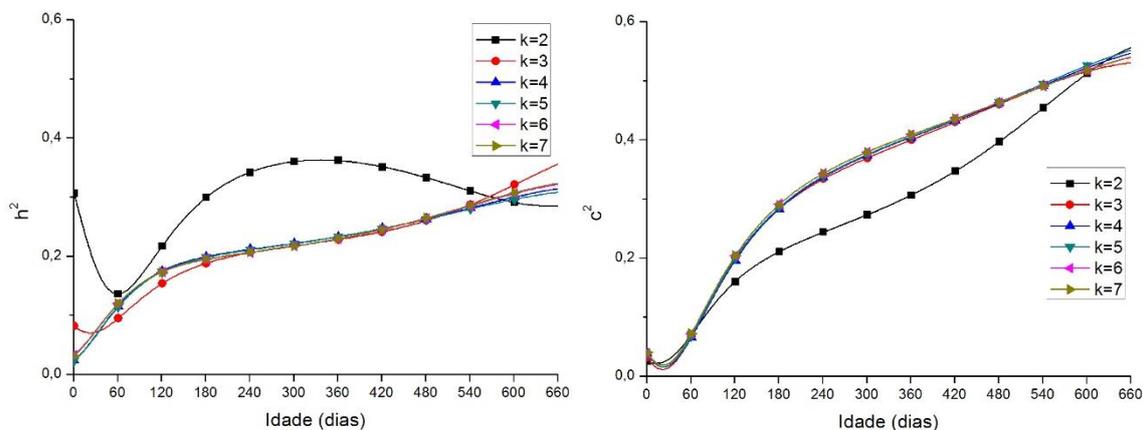


Figura 1. 3. Estimativas de herdabilidade direta (h^2) e de componente de variância para ambiente permanente animal como proporção da variância fenotípica (c^2), obtidas por meio de modelos que empregaram funções de diferentes ordens para representar a trajetória média.

As estimativas de h^2 e c^2 foram bem semelhantes entre as ordens analisadas, exceto pelo ajuste linear que se distinguiu dos demais, principalmente no meio da curva.

Para h^2 , dentre os modelos que se apresentaram semelhantes, o modelo quadrático ($k=3$) diferenciou-se no início e no final da curva. Resultados semelhantes foram relatados na literatura (SOUSA et al., 2010; SARMENTO, et al., 2010) ao comparar diferentes ordens para a modelagem da curva média de crescimento, em que, a herdabilidade direta de $k=2$ se mostrou superior às demais funções na maior parte da trajetória.

As estimativas de variâncias de ambiente permanente direto aumentaram à medida que o animal cresceu e praticamente não houve diferença nas trajetórias obtidas pelos diferentes modelos. Meyer (2001) descreveu padrão semelhante para pesos do nascimento à desmama de animais Hereford.

Nesse estudo, a fase de maior oscilação da herdabilidade foi no início da curva, essa variação pode ser explicada pela existência da heterogeneidade da variância residual, a qual não foi considerada nos modelos. Em um modelo que é considerada apenas uma classe de variância residual (homogênea) para toda a trajetória de peso, este valor tende a ser superestimado no início da curva, pois a tendência da variância residual real é que seja menor no início da trajetória.

Os parâmetros genéticos estimados são oriundos da partição da variância total, portanto, diferentes estruturas de variância residual podem influenciar a proporção da variância dos efeitos aleatórios considerados. Por esse motivo, as comparações dos parâmetros estimados ficam restritas aos modelos com diferentes ordens da curva fixa consideradas no estudo.

As estimativas de herdabilidade direta materna e de ambiente permanente materno, sob diferentes ordens de polinômios para ajuste da regressão fixa, estão representadas na figura 1.4.

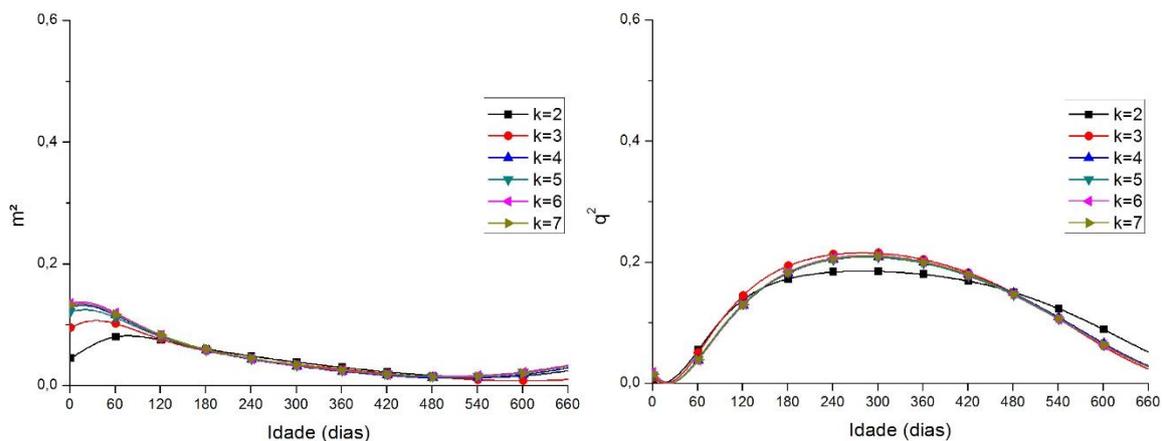


Figura 1. 4. Estimativas de herdabilidade materna (m^2) e de componente de variância para ambiente permanente materno como proporção da variância fenotípica (q^2) obtidas por meio de modelos que empregaram funções de diferentes ordens para representar a trajetória média.

Para o efeito genético materno, no início da curva, diferenciaram-se os ajustes linear e quadrático dos demais, sendo que após 120 dias, a semelhança entre os modelos fica mais evidente.

Após o nascimento, as estimativas da variância de ambiente permanente materna como proporção da variância fenotípica, aumentaram até próximo a desmama, diminuindo posteriormente, até os 660 dias. Essa estimativa indica a influência do ambiente materno sobre a característica de crescimento, sendo mais notável no período correspondente ao nascimento e à idade próxima aos 240 dias de idade.

As correlações para os valores genéticos dos animais considerando as diferentes ordens (k) de 2 a 7, preditos aos 240 e aos 660 dias de idade estão apresentadas na Tabela 1.3.

Tabela 1. 3. Correlação de Spearman entre os valores genéticos preditos para os pesos à desmama (acima da diagonal) e aos 660 dias de idade (abaixo da diagonal) por meio de modelos que empregaram funções de diferentes ordens para representar a trajetória média da po

Ordens	K=2	K=3	K=4	K=5	K=6	K=7
K=2	1,00	0,89	0,79	0,79	0,80	0,80
K=3	0,89	1,00	0,96	0,97	0,97	0,97
K=4	0,95	0,96	1,00	1,00	1,00	1,00
K=5	0,94	0,95	1,00	1,00	1,00	1,00
K=6	0,94	0,96	1,00	1,00	1,00	1,00
K=7	0,94	0,96	1,00	1,00	1,00	1,00

Observou-se que em ambas as idades, a correlação de Spearman é máxima (100%) entre os valores genéticos preditos a partir da função cúbica (ordem 4), ou seja, o ranqueamento dos animais de acordo com os valores genéticos preditos não é alterado quando considerada ordens de ajuste de 4 a 7 para a curva de regressão fixa.

Conclusões

O ajuste da regressão fixa apresentou influência sobre as estimativas dos parâmetros genéticos, indicando a necessidade de se modelar adequadamente a curva média de crescimento de bovinos de corte, por meio de modelos de regressão aleatória.

Uma função polinomial de ordem quatro (cúbica) pode ser utilizada para descrever a curva média de crescimento de bovinos da raça Nelore Mocho, criados na região Norte do Brasil, do nascimento aos 660 dias de idade.

Referências Bibliográficas

ALI, T. E.; SCHAEFFER, R. Accounting for covariances among test day milk yields in dairy cows. **Canadian Journal of Animal Science**, v.67, p.637-644, 1987

ARANGO, J.A.; CUNDIFF, L.V.; VAN VLECK, L. D. Covariance functions and random regression models for cow weight in beef cattle. **Journal of Animal Science**, v.82, p.54-67, 2004.

BALDI F.; MELLO, M.A.; ALBUQUERQUE L.G. Random regression analyses using B-splines functions to model growth from birth to adult age in Canchim cattle. **Journal of Animal Breeding and Genetics-Zeitschrift Für Tierzucht Und Zuchtungsbiologie**, v.: 127, p.: 433 - 441, 2010.

COBUCL, J. A. et al. Genetic parameters for milk production by using random regression models with different alternatives of fixed regression modeling. **Revista Brasileira de Zootecnia**, v. 40, n. 3, p. 557-567, 2011.

BOLIGON, A. A. et al. Random regression analyses using B-spline functions to model growth of Nelore cattle. **Animal**, v. 6, n. 02, p. 212-220, 2012.

KIRKPATRICK, M., LOFSVOLD, D., BULMER, M. Analysis of the inheritance, selection and evolution of growth trajectories. **Genetics**, v.124, p.979-993, 1990

MEYER, K. Modeling repeated records: covariance functions and random regression models to analyze animal breeding data. In: WORLD CONGRESS ON GENETICS APPLIED TO LIVESTOCK PRODUCTION, 6. 1998, Armindale, Australia. **Proceedings...** Armindale, Australia, v.25, p.517-520, 1998.

MEYER, K. Estimates of genetic and phenotypic covariance functions for postweaning growth and mature weight of beef cows. **Journal of Animal Breeding and Genetics**, v.116, n.3, p.181-203, 1999

MEYER, K. Estimates of genetic covariance functions for growth of Australian Angus cattle from random regression models fitting different orders of polynomials.” In: ANNUAL MEETING OF THE EUROPEAN ASSOCIATION FOR ANIMAL PRODUCTION, 54. 2003, Roma. **Proceedings...** Roma: EAAP, 2003. CDROM

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

SAKAGUTI, E.S.; SILVA, M.A.; MARTINS, E.N. et al. Trajetória de crescimento e efeito da idade da vaca nos modelos de regressão aleatória de bovinos da raça Tabapuã. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.54, n.4, p.414-423, 2002

SANTORO, K.R.; BARBOSA, S.B.P.; SANTOS, E. de S.; BRASIL, L.H. de A. Uso de funções de covariância na descrição do crescimento de bovinos Nelore criados no Estado de Pernambuco. **Revista Brasileira de Zootecnia**, v. 34, n. 6, p. 2290-2297, 2005.

SARMENTO, J.L.R.; REGAZZI, A.J.; SOUSA, W.H. et al. Modelagem da curva média de crescimento de ovinos Santa Inês em modelos de regressão aleatória. **Revista Brasileira de Zootecnia**, v. 39, 2010.

SOUSA, J.E.R.; SILVA, M.A.; SARMENTO, J.L.R.; SOUSA, W.H.; SOUZA, M.S.M. Avaliação da trajetória média de crescimento de caprinos em modelos de regressão aleatória. **Archivos de Zootecnia**, v. 59, n. 226, p. 267-276, 2010.

5 CAPÍTULO II

Modelos de regressão aleatória sob diferentes estruturas de variância residual para descrever o crescimento de bovinos Nelore Mocho

Resumo: O objetivo deste trabalho foi comparar modelos de regressão aleatória com diferentes estruturas de variância residual, na estimação dos componentes de covariância e parâmetros genéticos de características de crescimento. Utilizaram-se 15.148 registros de pesos de 3.115 bovinos da raça Nelore Mocho com idade entre 1 e 660 dias, criados na região Norte do Brasil e nascidos entre os anos de 1995 e 2010. Polinômios ortogonais de Legendre de ordem quatro foi usada para modelar a trajetória média de crescimento animal, e de ordem três, para modelar os efeitos genéticos aditivos direto e materno; e de ambiente permanente direto e materno. Diferentes estruturas de variâncias residuais foram consideradas, variando de um a sete o número de classes de variâncias residuais. Os modelos foram comparados pelo teste de razão de verossimilhança, o critério de informação de Akaike e o critério bayesiano de Schwarz. As estimativas de componentes de variância e parâmetros genéticos foram afetadas pela modelagem da variância residual. A utilização da variância residual homogênea não foi adequada para modelar a curva de crescimento, sendo constatada a necessidade da utilização de heterogeneidade de variâncias residuais. O modelo que considerou 6 classes de variâncias residuais foi considerado o mais adequado para a modelagem do crescimento dos bovinos Nelore Mocho, com idade entre 1 e 660 dias, criados na região Norte do Brasil.

Palavras-chave: modelagem de crescimento, homogeneidade de variância, resíduo, heterogeneidade de variância, Legendre.

5 CHAPTER II

Random regression models under different structures of residual variance to describe the growth of Nelore Mocho cattle

Abstract: The objective of this work was to compare random regression models with different residual variance structures in the estimation of covariance components and genetic parameters of growth characteristics. We used 15,148 weight records of 3,115 Nelore Mocho cattle aged between 1 and 660 days raised in northern region of Brazil, born between 1995 and 2010. Legendre orthogonal polynomials of order four were used to model the average trajectory of animal growth, and order three, to model the genetic additive and permanent environment effects, direct and maternal.. Different structures of residual variances were considered, ranging from one to seven, the number of classes residual variances. The models were compared by the likelihood ratio test, the Akaike information criterion and the Bayesian Schwarz criterion. Estimates of variance components and genetic parameters were affected by the modeling of residual variance. The use of the homogeneous residual variance was not adequate to model the growth curve, and it was verified the need to use heterogeneity of residual variances. The model that considered 6 classes of residual variances was considered the most suitable for the modeling of the Nelore Mocho cattle growth, aged between 1 and 660 days, raised in the Northern region of Brazil.

Keywords: Modeling of growth, homogeneity of variance, residue, heterogeneity of variance, Legendre.

Introdução

O melhoramento genético é uma ferramenta essencial para incrementar a produção de carnes no Brasil, principalmente a carne bovina, considerada a mais produzida e exportada pelo país. Dentre as diversas metodologias disponíveis para a avaliação genética de animais de produção, a utilização de modelos de regressão aleatória na avaliação da curva de crescimento é eficiente pela possibilidade de utilização de todas as informações de pesagens dos indivíduos, sem a necessidade de realizar pré-ajustes (Malhado et al., 2008).

Para o ajuste dos modelos de regressão aleatória é necessário o uso de uma matriz de covariâncias, sendo as funções de covariâncias uma alternativa eficiente (Sousa Júnior et al., 2010), além disso, o emprego destes modelos permite a modelagem da variância residual por meio de diferentes estruturas.

A necessidade de considerar heterogeneidade de variâncias residuais em modelos de regressão aleatória pode ser encontrada em vários trabalhos (El faro & Albuquerque, 2003, em bovinos de leite; Meyer, 2000; Sousa Júnior et al., 2010, em bovinos de corte). Para características de crescimento, as variâncias residuais tendem a aumentar à medida que o animal fica mais pesado, portanto, a modelagem da curva de crescimento considerando a variância de várias classes residuais em função da idade (heterogênea) pode melhorar a partição da variância fenotípica nas variâncias devido aos efeitos aleatórios incluídos no modelo de análise (Meyer, 2001; Fischer et al., 2004). Em contrapartida, o incremento no número de classes de variância residual deve ser ponderado, pois o excesso de classes no modelo implica no aumento do número de parâmetros a serem estimados, o que pode dificultar a análise.

O objetivo deste trabalho foi comparar os modelos de regressão aleatória com diferentes estruturas de variâncias residuais, bem como avaliar a influência da homogeneidade e heterogeneidade da variância residual na estimação dos componentes de covariância e parâmetros genéticos de características de crescimento de bovinos da raça Nelore Mocho, criados na região norte do Brasil.

Material e Métodos

Descrição dos dados

Os dados são provenientes de rebanhos participantes do Programa de Desenvolvimento Ponderal da Associação Brasileira de Criadores de Zebu (ABCZ). Estes rebanhos estão localizados em sete estados da região norte do Brasil (Amazonas, Acre, Amapá, Roraima, Rondônia, Pará e Tocantins).

Inicialmente, no banco de dados constava 44.106 registros de animais. A edição dos dados consistiu em eliminar: registros de pesos com idade acima de 660 dias; animais sem identificação do pai e/ou da mãe; animais com menos de 3 pesagens; e grupos contemporâneos com menos de 3 animais.

A definição de grupo contemporâneo incluiu: sexo, ano de nascimento, estação do ano ao nascimento, ano da pesagem, mês da pesagem, fazenda e condição de criação. As estações de nascimento foram classificadas em chuvosa (dezembro a maio) e seca (junho a novembro). As condições de criação foram: “mamando” e “desmamado”, com regime alimentar à pasto.

Uma análise de consistência dos dados foi feita com finalidade de excluir registros com possíveis erros de medição (*outliers*), através da constatação de uma variabilidade expressiva dentro de cada grupo contemporâneo. Foram excluídos os registros de peso que apresentavam valores fora do intervalo de ± 3 desvios-padrão em relação à média do grupo.

Após edição dos dados, havia 15.148 observações de peso pertencentes a 3.115 animais nascidos entre os anos de 1995 e 2010, criados em regime alimentar à pasto, com peso médio ao nascer de 28 kg. O peso médio e o número de registros por idade estão representados na figura 2.1.

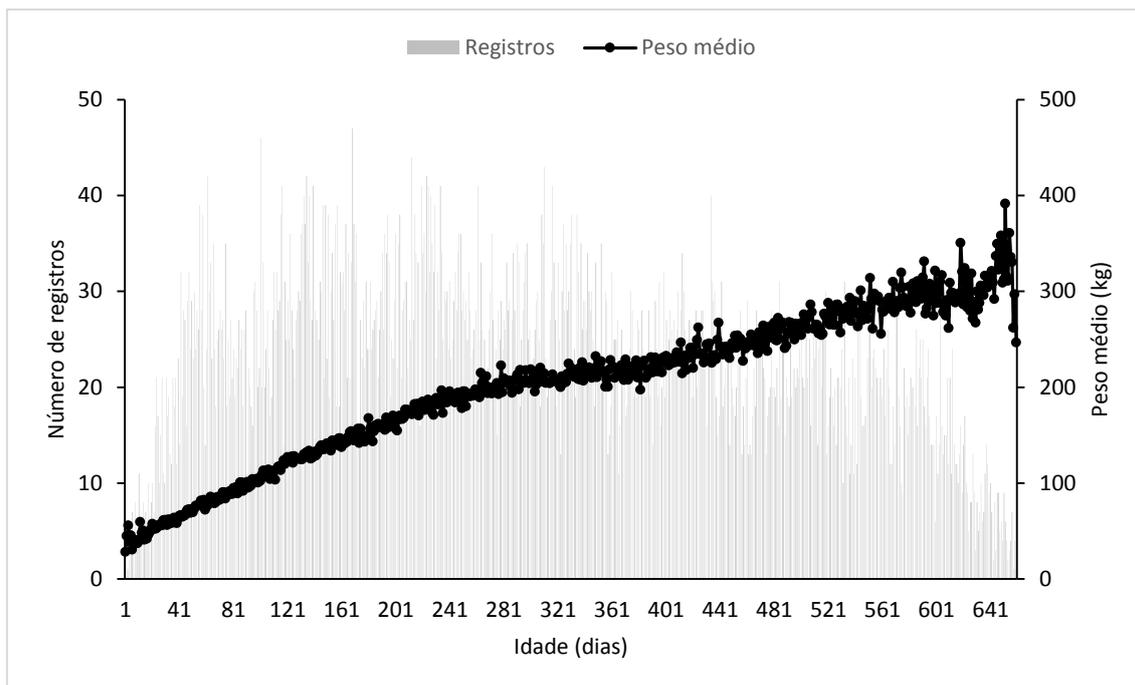


Figura 2. 1. Número de registro e peso corporal médio por idade de bovinos Nelores provenientes da região Norte do Brasil.

A estrutura de dos com número de observações por animal e informações referentes ao pedigree está representado na tabela 2.1.

Tabela 2.1.-Resumo da estrutura dos dados de bovinos Nelores provenientes da região Norte do Brasil utilizados no estudo

Informação	Quantidade
Dados	15.148
Animais com observação	3.115
Com 3 observações	712
Com 4 observações	698
Com 5 observações	549
Com 6 observações	659
Com 7 a 10 observações	497
Reprodutores	282
Matrizes	764
Animais na matriz de parentesco	5.678
Grupos contemporâneos	1.819

A edição dos dados e as análises de consistência foram feitas com auxílio do software SAS, versão 9.0 (2003).

Modelos e análises

Avaliou-se a modelagem da variância residual considerando-a homogênea e heterogênea por meio de diferentes estruturas, as quais contemplaram as idades-padrão, importantes para avaliação do desenvolvimento corporal (1, 240, 365, 550 dias de idade). Inicialmente, considerou-se classes de idade (CL_m , sendo m o número de classes), agrupadas como segue: CL1, homogênea; CL2, 1-240 e 241-660 dias; CL3, 1-240, 241-550 e 551-660 dias; CL4, 1-240, 241-365, 366-550 e 551-660 dias; CL5, 1-240, 241-330, 331-365, 366-550 e 551-660 dias; CL6, 1-120, 121-150, 151-240, 241-365, 366-550 e 551-660 dias; CL7, , 1-120, 121-150, 151-240, 241-365, 366-450, 451-550 e 551-660 dias; e CL8, 1-120, 121-150, 151-240, 241-365, 366-410, 411-450, 451-550 e 551 a 660 dias de idade.

Posteriormente, ajustaram-se diferentes modelos animais de regressão aleatória com ordem cúbica, para o ajuste da regressão fixa, e os efeitos aleatórios da regressão (genéticos e ambientais) com a ordem de ajuste quadrática.

Em notação matricial, o modelo acima e suas respectivas pressuposições podem ser descritos como:

$$Y = Xb + Z1a + Z2c + Z3m + Z4q + e$$

$$E \begin{bmatrix} Y \\ a \\ m \\ c \\ q \\ e \end{bmatrix} = \begin{bmatrix} Xb \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, e$$

$$\begin{aligned} \text{Var}(a) &= K_a \otimes A, \\ \text{Var}(m) &= K_m \otimes A, \\ \text{Var}(c) &= K_c \otimes I_{Nd}, \\ \text{Var}(q) &= K_q \otimes I_{Nd}, \\ \text{Var}(e) &= R; \end{aligned}$$

Em que, Y é um vetor de N observações referentes a Nd animais; b é um vetor que contém os efeitos fixos e os coeficientes da regressão fixa; a é um vetor $K_a \times N_D$ de coeficientes de regressão aleatória genéticos aditivos diretos, em que $N_D > Nd$ denota o número total de animais na análise, ou seja, na matriz de parentesco (5.678); c é um vetor $K_c \times N_d$ de coeficientes de regressão aleatória ambiente permanente de animal; m é um vetor $K_m \times N_D$ de coeficientes de regressão aleatória genéticos aditivos maternos; q é um vetor $K_q \times N_m$ de coeficientes de regressão aleatória de ambiente permanente materno, sendo N_m igual ao número de fêmeas que têm progênies com dados; e é um vetor de erros aleatórios; X , Z_1 , Z_2 , Z_3 e Z_4 referem-se às matrizes de incidência dos efeitos fixos, genético aditivo direto, ambiente permanente de animal, genético aditivo materno e ambiente permanente materno, respectivamente. K_a , K_c , K_m e K_q são as matrizes de (co)variâncias entre os coeficientes de regressão aleatórios genético

aditivo direto, ambiente permanente de animal, genético aditivo materno e ambiente permanente materno, respectivamente; A é a matriz dos numeradores do coeficiente de parentesco entre os indivíduos; I_{Nd} é uma matriz identidade de dimensão N_d ; I_{Nm} é uma matriz identidade de dimensão N_m ; \otimes é o produto de Kronecker; R é uma matriz diagonal de variâncias residuais, e sua dimensão depende da estrutura empregada na análise.

A covariância entre os efeitos genéticos aditivos direto e materno foi assumida como sendo igual a zero. As (co)variâncias entre os coeficientes de regressão aleatória genéticos aditivos direto e materno e ambiente permanente de animal e materno, conforme o modelo ajustado, foram estimadas pelo método da máxima verossimilhança restrita (REML), usando o software WOMBAT (2013).

Os diferentes modelos foram comparados, inicialmente, pela mudança no logaritmo da função de máxima verossimilhança (LogL), por meio do teste da razão de verossimilhança (TRV). A estatística TRV foi obtida pela seguinte expressão: $TRV_{ij} = 2 \text{ Log } L_i - 2 \text{ Log } L_j$, em que $\text{Log } L_i$ é o máximo da função de verossimilhança para o modelo completo i , e $\text{Log } L_j$ é o máximo da função de verossimilhança para o modelo reduzido j . A estimativa TRV foi comparada com o valor do qui-quadrado tabelado, com d graus de liberdade e nível de significância de 1%, em que d é a diferença entre o número de parâmetros estimados pelos modelos completo e reduzido. Foram usados, também, os critérios de informação de Akaike (AIC) e bayesiano de Schwarz (BIC), que permitem comparações entre modelos não aninhados e impõem penalidades, de acordo com o número de parâmetros estimados.

Posteriormente, avaliaram-se as mudanças ocorridas nos componentes de variância e parâmetros genéticos estimados ao longo da curva de crescimento em função do ajuste da variância residual, como também, na classificação dos animais (correlação de Spearman), com base nos valores genéticos preditos ao nascimento, à desmama (240 dias) e aos 660 dias de idade.

Resultados e Discussão

Ao se analisar inicialmente a modelagem do resíduo, por meio do Log L, observou-se melhoria no ajuste com o aumento do número de classes, porém, a significância desse ajuste só ocorre a partir do segundo modelo (Tabela 2.2).

Tabela 2. 2. Número de parâmetros (NP), Logaritmo da função de máxima verossimilhança (LogL), critério de informação de Akaike (AIC), critério de informação

Bayesiano (BIC) e teste da razão de verossimilhança (TRV) estimados com os modelos que empregaram diferentes

Modelo	NP	Log L	AIC	BIC	TRV
CL1	25	-43848,1	87746,2	87933,6	
CL2	26	-43847,6	87747,3	87942,2	(CL2-CL1) = 0,9 ^{ns}
CL3	27	-43842,7	87739,4	87941,8	(CL3-CL2) = 9,9 ^{**}
CL4	28	-43829,7	87715,3	87925,2	(CL4-CL3) = 26,0 ^{**}
CL5	29	-43828,5	87715,0	87932,4	(CL5-CL4) = 2,3 ^{ns}
CL6	30	-43808,4	87676,9	87901,8	(CL6-CL5) = 40,1 ^{**}
CL7	31	-43807,9	87677,9	87910,3	(CL7-CL6) = 1,0 ^{ns}
CL8	32	-43807,9	87679,9	87919,8	(CL8-CL7) = 0 ^{ns}

CL(número de classes); NS: Não-Significativo; **: Significativo a 1% de probabilidade.

A melhoria do ajuste ocorre gradualmente com o aumento do número de classes no modelo, sendo que, a maior diferença significativa ocorre entre os modelos CL6 e CL5.

Essas mudanças são compatíveis quando se compara os modelos por meio dos critérios AIC e BIC. O melhor resultado, quanto ao ajuste da variância residual, foi observado com a utilização do modelo CL6.

De acordo com os critérios Log L e TRV, utilizados para avaliar a qualidade do ajuste, o modelo menos adequado foi o que considerou homogeneidade de variância residual (CL1).

Ao comparar a classe homogênea (CL1) com as demais classes heterogêneas que apresentaram diferenças significativas (CL3, CL4 e CL6), percebe-se que a variância residual no início e no final da curva tendem a ser menor que no meio da trajetória (Figura 2.2). Portanto, o modelo de ajuste considerando apenas uma classe de variância (homogêneo) é inadequado, pois ignora as oscilações da variância residual, principalmente nos extremos.

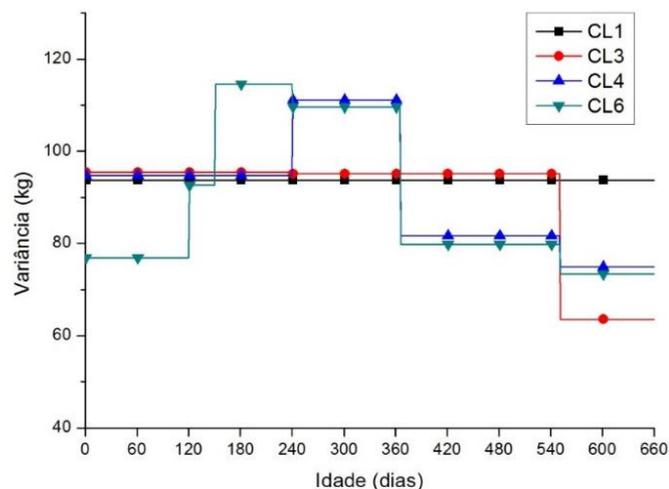


Figura 2. 2. Estimativas de variâncias residuais homogênea (CL1) e heterogêneas por meio das classes CL3, CL4 e CL6.

Na Figura 2.2, podemos perceber que as estimativas de variâncias residuais no início da trajetória de crescimento foram altas para todos os modelos, sendo menor para o CL6, que apresentou variância residual de 76,9. Toral et al. (2009), ao ajustar as variâncias residuais em bovinos de corte por intermédio de diferentes estruturas, relataram resultados semelhantes, em que a variância residual aumenta acentuadamente no início da curva e depois tende a decrescer.

Quando se admite homogeneidade de variância residual (Figura 2.2) na curva de crescimento, criam-se distorções consideráveis na partição da variância total (Sousa et al., 2008); portanto, houve implicações nas variâncias residuais estimadas, comparadas às estimativas obtidas quando se considerou a heterogeneidade de variâncias.

Na Figura 2.3 observa-se a implicação das diferentes classes residuais para as herdabilidades direta e materna e para a proporção da variância de ambiente permanente de ambiente materno para a variância fenotípica.

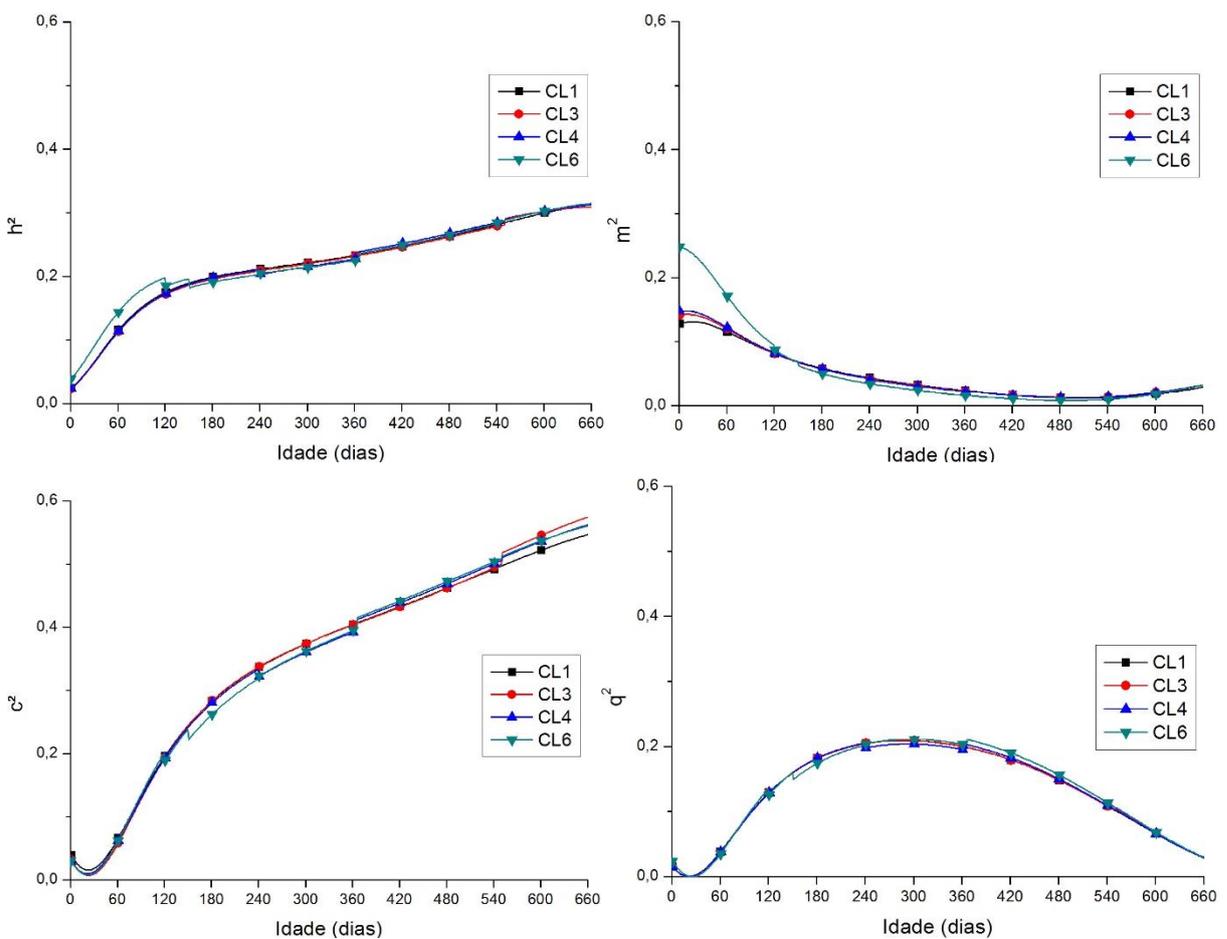


Figura 2. 3. Estimativas de herdabilidade direta (h^2 - acima à esquerda) e materna (m^2 - acima à direita) e componentes de variância para ambiente permanente de animal (c^2 - abaixo à esquerda) e materno (q^2 - abaixo à direita) como proporção da variância fenotípica

As estimativas de herdabilidade diretas e materna, obtidas ao longo da curva de crescimento, diferiram em função da estrutura de variâncias residuais ajustada, particularmente no início da curva.

A herdabilidade materna estimada pelo modelo com homogeneidade saíram de próximo a 0,12 próximo ao nascimento, para 0,25 com o modelo heterogêneo CL6. Estimativa de herdabilidades maternas diferentes de zero ao nascimento é biologicamente coerente, pois a mãe exerce influência sobre a cria desde o período de gestação, ou seja, a contribuição do efeito materno não começa a partir do nascimento do animal (Sarmiento et al., 2006).

O ambiente permanente de animal como proporção da variância fenotípica estimado com homogeneidade e heterogeneidade de variância residual, diferiu apenas da terça parte da curva até o final, sendo a magnitude dessa variação, bem baixa (Figura 2.3). Da mesma forma, o ambiente permanente de materno como proporção da variância fenotípica, praticamente não diferiu, durante o período estudado. Resultados diferentes foram observados por Sarmiento et al. (2006) e Sousa et al (2008) nos quais, as variações das estimativas entre os modelos foram bem mais expressivas.

Para observar a magnitude das diferenças entre os modelos homogêneo (CL1) e o modelo heterogêneo (CL6), observa-se os valores genéticos preditos para os cinco melhores animais nas idades: nascimento, 240 e 660 dias (Tabela 2.3).

Tabela 2. 3. Valores genéticos preditos para os cinco melhores animais para pesos ao nascimento, aos 240 e 660 dias de idade considerando homogeneidade e heterogeneidade (CL6) de variância residual.

Idade (dias)	Animal	Valor Genético Predito (kg)			
		Homogêneo	CL6	Desvio (%) ¹	Correlação (ρ)
Nascimento	18358	2,26	3,08	-26,61	0,97
	1291	2,35	3,22	-27,01	
	18367	2,61	3,42	-23,80	
	10755	3,29	4,60	-28,40	
	12032	3,12	4,61	-32,48	
240 dias	17206	20,47	20,92	-2,17	0,99
	10755	20,54	21,16	-2,91	
	1045	21,26	21,18	0,36	
	17917	22,35	21,70	2,98	
	17615	23,23	23,13	0,44	
660 dias	13617	51,34	52,62	-2,42	0,99
	17206	55,77	57,23	-2,56	
	17536	63,51	64,38	-1,37	
	17615	64,19	64,90	-1,10	
	1045	66,50	67,35	-1,26	

¹Os desvios dos valores genéticos preditos pelos modelos homogêneo e CL6, expressos em porcentagem.

Observa-se na Tabela 2.3 que os valores genéticos preditos assumindo homogeneidade foram quase todos subestimados em relação aos preditos quando a variância residual foi heterogênea.

Analisando os desvios dos valores genéticos preditos, em percentagem, os maiores desvios foram observados no peso ao nascer, variando de 23,80 a 32,48% para os cinco melhores animais da população.

Constatou-se também, que a correlação de Spearman entre os modelos para a idade ao nascer foi de 0,97. Portanto, apesar da diferença do valor genético predito, entre os modelos homogêneo e CL6, ser de média magnitude na idade ao nascer (Tabela 2.3), a correlação de Spearman revela que a variação no rank dos melhores animais foi pequena. Para as idades 240 dias e 660 dias, a variação relativa foi baixa, em torno de 1 a 3% do valor genético.

Na Figura 2.4 observamos as correlações para os valores genéticos preditos em diferentes idades obtidas para o modelo mais adequado (CL6).

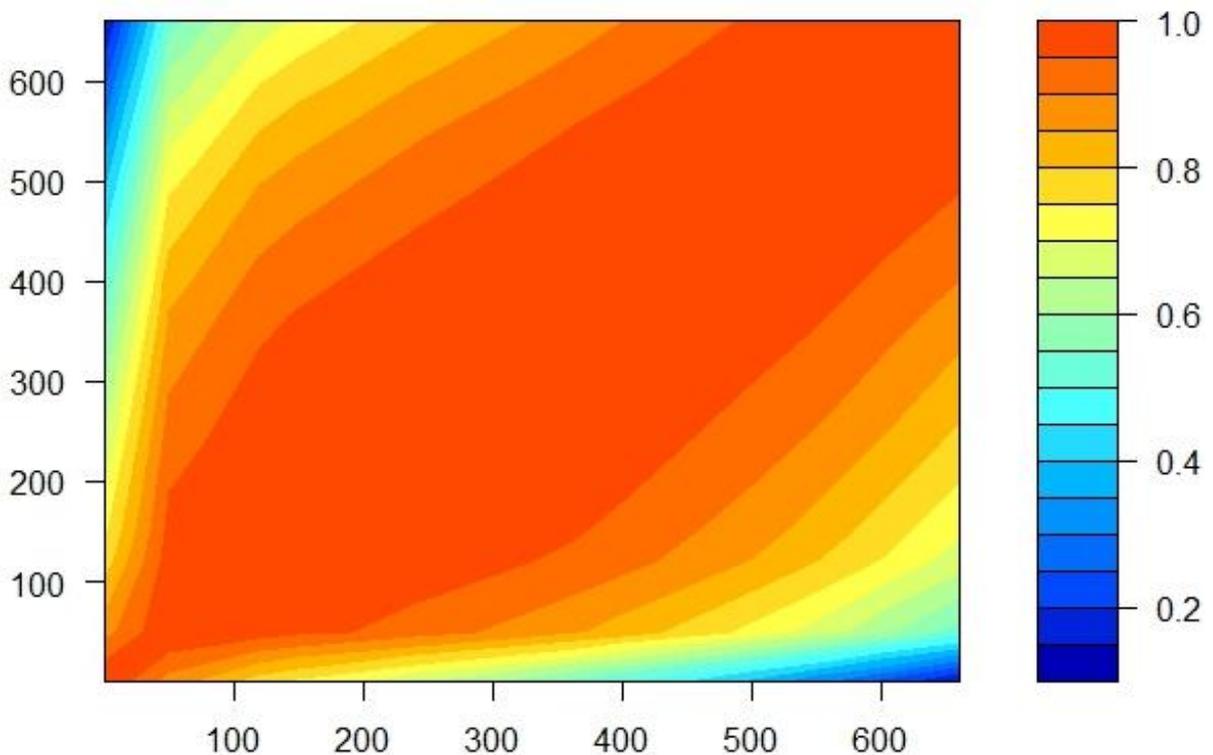


Figura 2. 4. Correlação de Spearman para o valor genético aditivo direto, obtido em diferentes idades, por meio do modelo composto por 6 classes de variância residual (CL6).

Na Figura 2.4, observa-se que a magnitude da correlação entre os valores genéticos em diferentes idades é representada por um gradiente de cores, como indicado na legenda ao lado do gráfico.

Como esperado, a correlação entre os valores genéticos tende a aumentar, quando estes, são preditos em pontos de idades mais próximas entre si. Em contrapartida, a correlação tende a ser nula à medida que a distância entre as idades analisadas tende a ser a máxima (nesse estudo, entre 1 e 660 dias de idade).

A predição dos valores genéticos dos animais de forma precoce é um importante artifício para tomada de decisões no âmbito do melhoramento genético. Com base nos resultados da Figura 2.4, a correlação para os valores genéticos apresentou-se de moderada a alta entre a idade máxima (660 dias) e à desmama (em torno de 240 dias). Portanto, a seleção dos melhores animais para característica de peso aos 240 dias pressupõe a seleção dos melhores animais aos 660 dias de idade.

Conclusões

As estimativas de componentes de variância e parâmetros genéticos que caracterizam a curva de crescimento de bovinos da raça Nelore Mocho foram afetadas pela modelagem da variância residual. É necessária a utilização de heterogeneidade de variâncias residuais para modelar corretamente as variâncias associadas à curva de crescimento.

O modelo que considerou 6 classes de variância residual foi o mais adequado para a modelagem do crescimento dos bovinos Nelore Mocho, criados na região norte do Brasil, do nascimento aos 660 dias de idade.

Referências Bibliográficas

EL FARO, L.; ALBUQUERQUE, L. G. Predição de valores genéticos para a produção de leite no dia do controle e para produção acumulada até 305 dias. **Revista Brasileira de Zootecnia**, v.34, n.2, p.496-507, 2005.

FISCHER, T.M.; VAN DER WERF, J.H.J.; BANKS, R.G. et al. Description of lamb growth using random regression on field data. **Livestock Production Science**, v.89, p.175-185, 2004

MALHADO, C.H.M.; CARNEIRO, P.L.S.; PEREIRA, D.G.; et al. Progresso genético e estrutura populacional do rebanho Nelore no Estado da Bahia. **Pesquisa Agropecuária Brasileira**, v.43, p. 1163-1169, 2008.

MEYER, K. Random regression to model phenotypic variation in monthly weights of Australian beef cows. **Livestock Production Science**, v.65, p.19-38, 2000.

MEYER, K. Estimates of direct and maternal covariance functions for growth of Australian beef calves from birth to weaning. **Genetics Selection Evolution**, v.33, p.1-28, 2001.

SAKAGUTI, E.S.; SILVA, M.A.; QUAAS, R.L. et al. Avaliação do crescimento de bovinos jovens da raça Tabapuã, por meio de análises de funções de covariâncias. **Revista Brasileira de Zootecnia**, v.32, n.4, p.864-874, 2003.

SARMENTO, J.L.R.; REGAZZI, A.J.; SOUSA, W.H. et al. Estudo da curva de crescimento de ovinos Santa Inês. **Revista Brasileira de Zootecnia**, v.35, n.2, p.435-442, 2006.

SOUZA, J. C.; SILVA, L. O. C.; SIMÕES, G. H.; MOSER, J. T.; OSTAPECHEN, J.; PINTO, P. H. N.; RUVIERO, V.; MALHADO, C. H. M.; FERRAZ FILHO, P. B.; FREITAS, J. A.; SERENO, J. R. B. Tendências ambientais e genéticas para características produtivas de bovinos da raça Nelore. **Archivos Latinoamericanos de Producción Animal**, v. 16, n. 2, p. 85-90, 2008.

SOUSA JÚNIOR, S.C.; OLIVEIRA, S.M.P.; ALBUQUERQUE, L.G. et al. Estimação de funções de covariância para características de crescimento da raça Tabapuã utilizando modelos de regressão aleatória. **Revista Brasileira de Zootecnia**, v.39, p.1037-1045, 2010.

TORAL, F. B.; ALENCAR, M. A.; FREITAS, A. R. Estruturas de variância residual para estimação de funções de covariância para o peso de bovinos da raça Canchim. **Revista Brasileira de Zootecnia**, v.38, p.2152-2160, 2009.

6 CAPÍTULO III

Funções de covariâncias sob polinômios de Legendre para modelagem do crescimento de bovinos da raça Nelore Mocho da região Norte do Brasil

Resumo: Funções polinomiais da idade de diferentes ordens foram avaliadas na modelagem dos efeitos aleatórios associados ao estudo genético da curva de crescimento de bovinos. Utilizaram-se 15.148 registros de pesos de 3.115 de bovinos Nelore Mocho, com idade entre 1 e 660 dias, criados na região norte do Brasil e nascidos entre os anos de 1995 e 2010. Os efeitos fixos incluídos nos modelos de análise foram: grupos de contemporâneos, a idade da vaca ao parto (efeito linear e quadrático) e os polinômios de Legendre da idade de ordem quatro (cúbica) para representar a curva média de crescimento. Na parte aleatória, consideraram-se funções de diferentes ordens para modelar as variâncias associadas aos efeitos genéticos aditivos direto e materno e aos efeitos de ambiente permanente do animal e da mãe. A variância residual foi ajustada por meio de seis classes heterogêneas, distribuídas ao longo do período analisado. De acordo com o valor do critério AIC e BIC o modelo 6333 permitiu ajustar as mudanças das variâncias e covariâncias com o tempo, podendo ser utilizado para descrever as mudanças nas variâncias com a idade dos bovinos Nelore Mocho criados na região Norte do Brasil.

Palavras-chave: modelos de regressão aleatória, modelos lineares, trajetória de crescimento, parâmetros genéticos.

6 CHAPTER III

Covariance functions on Legendre polynomials for modeling the growth of Nelore Mocho cattle in the Northern region of Brazil

Abstract: Polynomial functions of the age of different orders were evaluated in the modeling of the random effects associated to the genetic study of the growth curve of cattle. We used 15,148 weight records of 3115 Nelore Mocho cattle aged between 1 and 660 days, raised in northern Brazil born between 1995 and 2010. The fixed effects included in the analysis models were: groups of contemporaries, cow's age at calving (linear and quadratic effect) and Legendre polynomials of age four (cubic) to represent the mean growth curve. In the random part, functions of different orders were considered to model the variances associated with the direct and maternal additive genetic effects and the permanent environment effects of the animal and the mother. The residual variance was adjusted by means of six heterogeneous classes, distributed over the analyzed period. According to the value of the AIC and BIC criterion, the model 6333 allowed to adjust the variance and covariance changes over time and could be used to describe the changes in the age variances of the Nelore Mocho cattle reared in the northern region of Brazil.

Keywords: Random regression models, linear models, growth trajectory, genetic parameters

Introdução

O melhoramento genético é uma ferramenta essencial para incrementar a produção de carnes no Brasil, principalmente a carne bovina, considerada a mais produzida e exportada pelo país. Dentre as diversas metodologias disponíveis para a avaliação genética de animais de produção, a utilização de modelos de regressão aleatória na avaliação da trajetória de crescimento é eficiente por utilizar todas as informações de pesagens dos indivíduos, sem necessidade de realizar ajustes de peso às idades desejadas (Malhado et al., 2008). Para o ajuste dos modelos de regressão aleatória é necessário o uso de uma matriz de covariâncias, sendo as funções de covariâncias uma alternativa eficiente na estimação desses componentes (Sousa Júnior et al., 2010).

Ao ajustar um modelo de regressão aleatória, implicitamente, ajusta-se uma função contínua que permite descrever as mudanças genéticas e ambientais com o passar do tempo. Vários trabalhos têm empregado como funções contínuas os polinômios ortogonais de Legendre (Kirkpatrick & Heckman, 1989; Meyer & Hill, 1997; Jamrozik et al., 1997; Meyer, 2007; Boligon et al., 2010), os quais podem requerer ordens de ajustes diferentes para cada efeito aleatório contido no modelo de análise.

Os programas de avaliação genética de bovinos de corte contemplam características de crescimento, por meio da seleção dos melhores animais segundo seus valores genéticos preditos, conseqüentemente, o sucesso da atividade depende da acurácia dos valores genéticos preditos pelos modelos de crescimento.

Nesse contexto, é imprescindível a utilização de modelos mais adequados para a obtenção de estimativas de parâmetros genéticos confiáveis, com o objetivo de classificar os melhores animais com potencial genético para incrementar a produção média da população (Boligon et al., 2010; Gonçalves, et al., 2011 e Laureano et al., 2011).

Do efetivo bovino brasileiro, a raça Nelore destaca-se com 90% de toda a população Zebuína do país (ABCZ, 2012). Portanto, neste capítulo objetivou-se analisar diferentes modelos de regressão aleatória com a finalidade de determinar o mais adequado para descrever mudanças nos parâmetros de avaliação do crescimento de bovinos da raça Nelore Mocho criados na região Norte do Brasil.

Material e Métodos

Descrição dos dados

Os dados são provenientes de rebanhos participantes do Programa de Desenvolvimento Ponderal da Associação Brasileira dos Criadores de Zebu (ABCZ). Estes rebanhos estão localizados nos sete estados da região norte do Brasil (Amazonas, Acre, Amapá, Roraima, Rondônia, Pará e Tocantins).

Inicialmente, no banco de dados constava com 44.106 registros de animais. A edição dos dados consistiu em eliminar: registros de pesos com idade acima de 660 dias; animais sem identificação do pai e/ou da mãe; animais com menos de 3 pesagens; e grupos contemporâneos com menos de 3 animais.

A definição de grupo contemporâneo incluiu: sexo, ano de nascimento, estação do ano ao nascimento, ano da pesagem, mês da pesagem, fazenda e condição de criação. As estações de nascimento foram classificadas em chuvosa (dezembro a maio) e seca (junho a novembro). As condições de criação foram: “mamando” e “desmamado”, com regime alimentar à pasto.

Uma análise de consistência dos dados foi feita com finalidade de excluir registros com possíveis erros de medição (*outliers*), através da constatação de uma variabilidade expressiva dentro de cada grupo contemporâneo. Foram excluídos os registros de peso que apresentavam valores fora do intervalo de ± 3 desvios-padrão em relação à média do grupo.

Após edição dos dados, havia 15.148 observações de peso pertencentes a 3.115 animais nascidos entre os anos de 1995 a 2010 e criados em regime alimentar à pasto, com peso médio ao nascer de 28 kg. O peso médio e o número de registros por idade estão representados na figura 3.1.

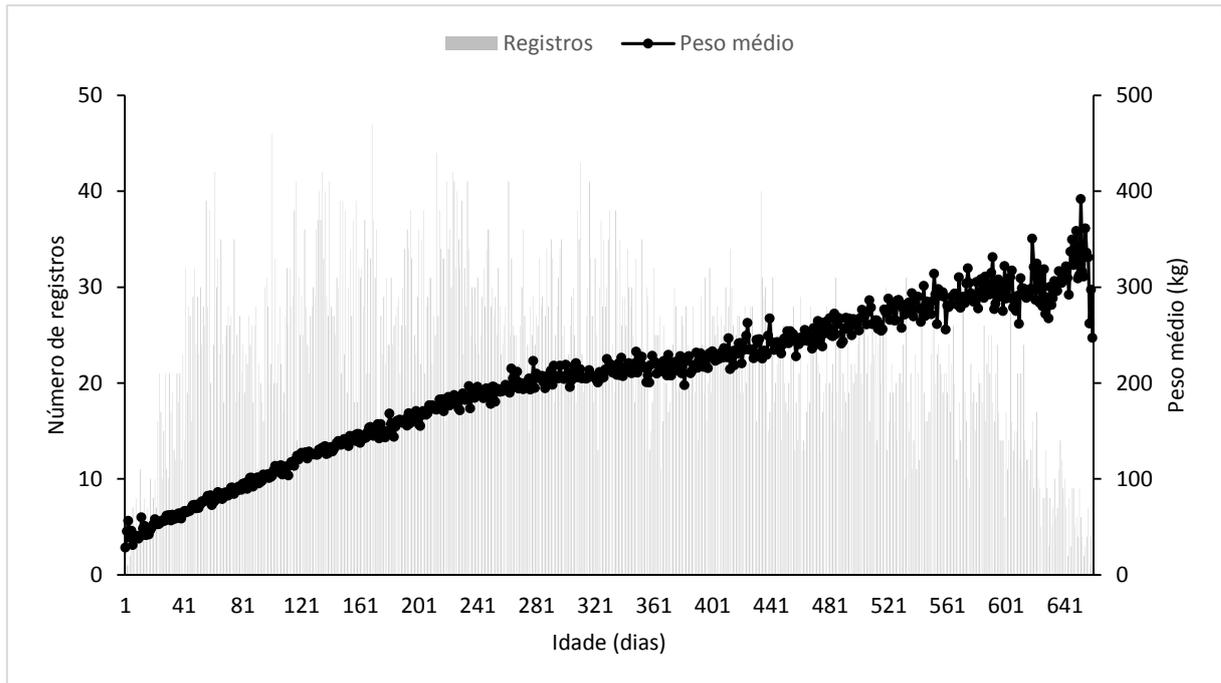


Figura 3. 1. Número de registro e peso corporal médio por idade de bovinos Nelores provenientes da região norte do Brasil.

Os pesos dos animais foram agrupados em 117 classes de idades. Cada classe foi composta a cada intervalo de 5 dias de idade, exceto no início e no final da curva, em que o intervalo de classes foi de 15 dias de idade, devido a diminuição na frequência dos registros nas idades localizadas nos extremos da curva, como se pode observar na Figura 3.1.

A estrutura de dados com número de observações por animal e informações referentes ao pedigree está representado na tabela 3.1.

Tabela 3. 1. Resumo da estrutura dos dados de bovinos Nelores provenientes da região norte do Brasil utilizados no estudo.

Informação	Quantidade
Dados	15.148
Animais com observação	3.115
Com 3 observações	712
Com 4 observações	698
Com 5 observações	549
Com 6 observações	659
Com 7 a 10 observações	497
Reprodutores	282
Matrizes	764
Animais na matriz de parentesco	5.678
Grupos contemporâneos	1.819

A edição dos dados e as análises de consistência foram feitas com auxílio do software SAS, versão 9.0 (2003).

Modelos e análises

Nos modelos de regressão aleatória, foram considerados na parte fixa, os efeitos de grupo de contemporâneo e, como covariáveis, a idade do animal à pesagem e idade da vaca ao parto (efeitos linear e quadrático). Como aleatórios, foram considerados os efeitos genético aditivo direto, genético aditivo materno, ambiente permanente animal e ambiente permanente materno. Um polinômio ortogonal de Legendre de ordem cúbica foi considerado o mais adequado para modelar a curva média da população.

O modelo geral pode ser representado por:

$$Y_{ij} = EF + \sum_{m=1}^4 \beta_m \Phi_m + \sum_{m=1}^{Ka-1} \alpha_{im} \Phi_m + \sum_{m=1}^{km-1} \gamma_{im} \Phi_m + \sum_{m=1}^{kc-1} \delta_{im} \Phi_m + \sum_{m=1}^{kq-1} \lambda_{im} \Phi_m + \varepsilon_{ij}$$

Em que Y_{ij} é o peso no $j^{\text{ésimo}}$ dia do $i^{\text{ésimo}}$ animal; EF refere-se a um conjunto de efeitos fixos, constituído pelo grupo de contemporâneo (1.819 subclasses) e pela covariável idade da vaca ao parto, efeitos linear e quadrático; β_m é o coeficiente de regressão fixo do peso sobre o polinômio de Legendre, representado por uma função cúbica, para modelar a curva média de crescimento da população; Φ_m é a função polinomial de Legendre da idade padronizada ($-1 < \text{idade} < 1$); $\alpha_{im}, \gamma_{im}, \delta_{im}, \lambda_{im}$ são os coeficientes de regressão genético aditivo direto, genético aditivo materno, ambiente permanente de animal, e ambiente permanente materno, respectivamente, para cada animal; Ka, km, kc, kq são ordem dos polinômios a serem ajustados para cada efeito aleatório; ε_{ij} = erro aleatório associado a cada idade i do animal j .

Na forma matricial, o modelo anterior, com suas respectivas pressuposições, pode ser escrito como:

$$Y = X\beta + Z_1\alpha + Z_2\gamma + W_1\delta + W_2\lambda + \varepsilon, \text{ sendo}$$

$$v \begin{bmatrix} \alpha \\ \gamma \\ \delta \\ \lambda \\ \varepsilon \end{bmatrix} = \begin{bmatrix} K_a \otimes A & 0 & 0 & 0 & 0 \\ 0 & K_m \otimes A & 0 & 0 & 0 \\ 0 & 0 & K_c \otimes I & 0 & 0 \\ 0 & 0 & 0 & K_q \otimes I & 0 \\ 0 & 0 & 0 & 0 & R \end{bmatrix}$$

Em que, Y é um vetor de N observações referentes a Nd animais; β é um vetor que contém os efeitos fixos e os coeficientes b_m da regressão fixa; α é um vetor $K_a \times N_D$ de coeficientes de regressão aleatória genéticos aditivos diretos, em que $N_D > Nd$ denota o número total de animais na análise, ou seja, na matriz de parentesco (5.678); γ é um vetor $K_m \times N_D$ de coeficientes de regressão aleatória genéticos aditivos maternos; δ é um vetor $K_c \times N_d$ de coeficientes de regressão aleatória ambiente permanente de animal; λ é um vetor $K_q \times N_m$ de coeficientes de regressão aleatória de ambiente permanente materno, sendo N_m igual ao número de fêmeas que têm progênie com dados; ε é um vetor de erros aleatórios; X , Z_1 , Z_2 , W_1 e W_2 referem-se às matrizes de incidência dos efeitos fixos, genético aditivo direto, genético aditivo materno, ambiente permanente de animal, e ambiente permanente materno, respectivamente.

K_a , K_c , K_m e K_q são as matrizes de (co)variâncias entre os coeficientes de regressão aleatórios genético aditivo direto, ambiente permanente de animal, genético aditivo materno e ambiente permanente materno, respectivamente; A é a matriz dos numeradores do coeficiente de parentesco entre os indivíduos; I_{Nd} é uma matriz identidade de dimensão N_d ; I_{N_m} é uma matriz identidade de dimensão N_m ; \otimes é o produto de Kronecker; R é uma matriz diagonal de variâncias residuais, e sua dimensão depende da estrutura empregada na análise;

A covariância entre os efeitos genéticos aditivos direto e materno foi assumida como sendo igual a zero. As (co)variâncias entre os coeficientes de regressão aleatória genéticos aditivos direto e materno e ambiente permanente de animal e materno, conforme o modelo

ajustado, foram estimadas pelo método da máxima verossimilhança restrita (REML), usando o software WOMBAT (Meyer, 2006) .

Os diferentes modelos foram comparados pelo logaritmo da função de verossimilhança (Log L); critério de informação de Akaike (AIC) e o critério de informação Bayesiano de Schwarz (BIC).

Resultados e Discussão

As médias dos pesos e os desvios-padrão aumentaram linearmente do nascimento até os 660 dias de idade (Tabela 3.2), com médias variando de $30,7 \pm 4,1$ a $347,5 \pm 121,7$ kg, respectivamente. Tais resultados assemelham-se aos descritos por Ferreira et al. (2015) para animais da raça Nelore.

Tabela 3. 2. Análise descritiva dos dados de peso à idade referentes aos bovinos Nelore Mocho criados na região norte do Brasil.

Idade (dias)	N	Média (Kg)	DP	CV (%)	Mínimo (Kg)	Máximo (kg)
1	93	30,7	4,1	13,3	28	56
120	77	124,1	20,6	16,6	89	172
240	116	189,4	37,2	19,6	111	370
365	77	242,7	50,3	20,7	155	415
550	68	290,6	85,0	29,3	180	674
660	8	347,5	121,7	35,0	182	520

N: Frequência; DP: desvio-padrão; CV: coeficiente de variação.

O emprego de modelos de regressão aleatória por meio de funções polinomiais de Legendre careceu de uma ordem básica (quadrática) para cada efeito aleatório considerado no modelo de análise. Posteriormente, as ordens de ajuste das funções contínuas foram aumentadas gradativamente, de modo a encontrar uma ordem mínima necessária para descrever as variâncias e a estrutura de covariância em função do tempo para cada efeito aleatório.

O resumo das análises utilizando modelos de regressão aleatória está descrito na Tabela 3.3, com os valores de Log L, AIC e BIC apresentados como desvio em relação ao menor valor.

Tabela 3. 3. Logaritmo da função de máxima verossimilhança (Log L), critério de informação de Akaike (AIC), critério de informação Bayesiano (BIC), número de parâmetros estimados (NP), expressos como desvio absoluto do menor valor.

Modelo ¹	NP	Log L ²	AIC ²	BIC ²
1- 2222	19	739.0	1422.3	1219.8
2- 3333	31	332.6	633.3	520.9
3- 5333	40	47.3	80.8	35.8
4- 5433	44	46.2	86.5	71.5
5- 5533	49	45.3	94.8	117.3
6- 5343	44	38.9	72.0	57.0
7- 5543	53	34.9	81.9	134.4
8- 6333	46	0.9	0.0	0.0
9- 6433	50	0.0	6.2	36.2

¹ Ordem de ajuste para os efeitos genético aditivo direto, genético aditivo materno, ambiente permanente de animal, e ambiente permanente da materno, respectivamente.

² Valores expressos como desvio do menor valor.

As ordens de ajuste das funções contínuas foram definidas sequencialmente, iniciando com o efeito genético direto e efeito genético materno, e posteriormente os outros efeitos. O incremento na ordem se deu por meio da observância dos valores dos critérios, procurando obter uma ordem que melhor se adeque para cada efeito.

Alguns autores como Lobo et al. (2000) e Albuquerque et al. (2004) estudando a modelagem da curva de crescimento em bovinos de corte, por meio de polinômios ortogonais de Legendre, afirmaram que os modelos parcimoniosos são os preferíveis, ou seja, deve-se dar preferência a modelos que envolvam o mínimo de parâmetros possíveis a serem estimados e que expliquem bem o comportamento da variável resposta. Outro aspecto, positivo é que, segundo Meyer (2005), haveria menor demanda computacional envolvida nas análises e consequente menor susceptibilidade a erros numéricos.

De acordo com o critério de Log L, as ordens mais altas para os efeitos aleatórios produziram melhores valores, porém, o aumento excessivo na ordem de ajuste dos efeitos contribuíram para uma maior parametrização do modelo. O modelo 9, com 50 parâmetros,

proporcionou melhor ajuste segundo o Log L, no entanto, este não deve ser o único critério a ser utilizado.

Ao observar os resultados obtidos com critério de informação de Akaike (AIC), verificou-se que houve melhoria nos seus valores com o aumento das ordens de ajustes até certo ponto. De acordo com este critério, o modelo considerando funções com as ordens 6333 (45 parâmetros) foi superior aos demais modelos no ajuste das (co)variâncias ao longo da trajetória de crescimento.

Corroborando com a afirmação anterior, o critério de informação Bayesiano (BIC), que impõe fatores de penalidades mais rigorosas que o AIC em função do número de parâmetros estimados, está de acordo com a indicação do modelo 8 (6333) para ajuste das (co)variâncias ao longo da trajetória de crescimento.

Dessa forma, o modelo 8 apresentou os menores valores de AIC e BIC, indicando que, embora não tenha tido o melhor valor de Log L, o modelo é eficiente para explicar adequadamente a variação dos dados, sem recorrer a um número excessivo de parâmetros.

Fischer et al. (2004) ajustaram modelos menos parametrizados, sendo o de maior ordem igual a três. Os autores verificaram que o melhor ajuste foi obtido com o modelo que empregou ordens igual a 3233, respectivamente, para os efeitos genéticos aditivos direto e materno e ambiente permanente de animal e da mãe.

De modo geral, os maiores valores de variância foram obtidos para o intercepto de todos os efeitos (Tabela 3.4). A correlação entre o intercepto e o coeficiente quadrático foram negativos para todos os efeitos aleatórios. Além disso, foi observada uma correlação positiva entre o intercepto e o coeficiente linear para todos os efeitos aleatórios.

Tabela 3. 4. Estimativas de variâncias (diagonal), covariâncias (abaixo da diagonal) e correlações (acima da diagonal) entre os coeficientes de regressão aleatória e autovalores associados à matriz dos coeficientes para o modelo com $k= 6333$.

	0	1	2	3	4	5	Autovalor
Efeito genético aditivo ($k_a=6$)							
0	228,88	0,72	-0,42	0,15	0,34	-0,10	73,09
1	97,47	80,69	-0,24	-0,57	0,27	-0,01	13,24
2	-26,31	-9,20	17,50	0,05	-0,84	0,15	7,16
3	9,80	-22,18	0,87	19,07	-0,10	-0,04	5,87
4	21,99	10,46	-15,02	-1,87	18,07	0,05	0,64
5	-7,04	-0,61	2,93	-0,88	1,11	22,77	0,00
Efeito genético aditivo materno ($k_m=3$)							
0	35,48	0,70	-0,44				91,13
1	6,76	2,61	0,33				8,86
2	-4,47	0,90	2,87				0,00
Efeito de ambiente permanente animal ($k_c=3$)							
0	367,24	0,88	-0,31				92,58
1	172,48	103,57	0,17				7,42
2	-26,01	7,43	18,94				0,00
Efeito de ambiente permanente materno ($k_q=3$)							
0	130,62	0,99	-0,99				99,99
1	29,74	6,77	-0,99				0,01
2	-37,42	-8,52	10,72				0,00

Em todos os efeitos aleatórios incluídos no modelo, o intercepto explicou a maior proporção da variância, sendo que no efeito de ambiente permanente materno, o intercepto explicou quase sua totalidade (99,99%).

Quando um modelo é muito parametrizado, é mais frequente que os autovalores apresentem valores próximos de zero, o que pode ter sido causado pelo aumento da multicolinearidade devido ao aumento da ordem de ajuste da função empregada (Sarmiento et al., 2010).

As estimativas de variância genética aditiva direta, de ambiente permanente animal e a variância fenotípica apresentaram tendência de crescimento à medida que aumentou a idade dos animais (Figura 3.2).

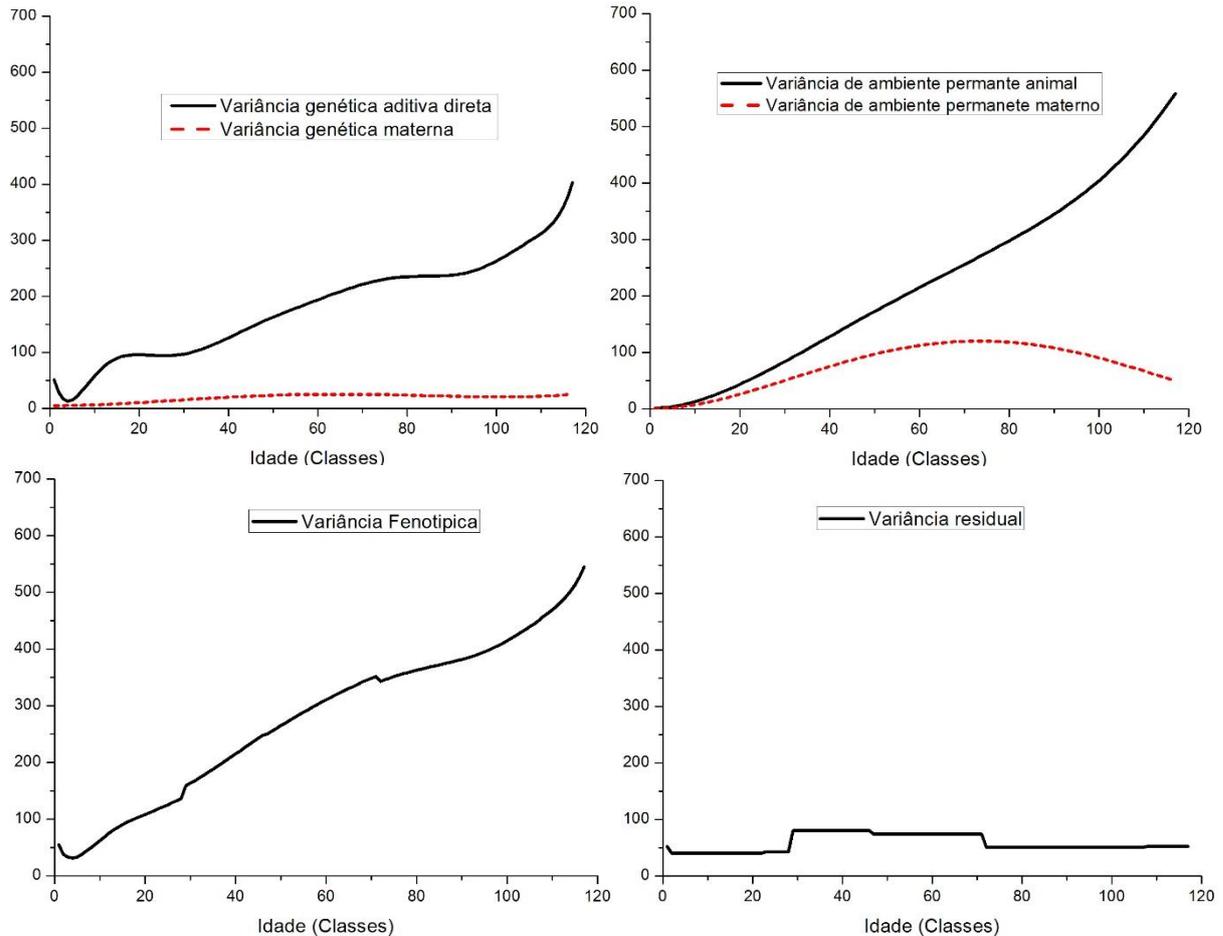


Figura 3. 2. Estimativas de componentes de variância genética aditiva direta e genética materna (acima à esquerda), de ambiente permanente de animal de ambiente permanente materno (acima à direita), metade da variância fenotípica (abaixo à esquerda) e variância residual (abaixo à direita).

Os maiores valores de variância de ambiente permanente direto refletiram na variância fenotípica, principalmente no final da curva. Albuquerque & Meyer (2001) e Boligon et al. (2010) relataram comportamentos similares para o efeito de ambiente permanente em avaliações de trajetórias de crescimento de bovinos da raça Nelore.

Para o efeito materno direto os resultados mostraram uma menor variância desses efeitos na participação da variância fenotípica total, em que, a variação apresentou um leve crescimento no início da trajetória e permaneceu praticamente constante após os 300 dias. No efeito de ambiente permanente materno, os resultados apresentaram um crescimento acentuado do nascimento até, aproximadamente, 400 dias de idade. Em geral, na maioria dos estudos envolvendo a inclusão dos efeitos maternos na análise apresentaram redução após o período da

desmama, indicando que tais efeitos têm pouca importância em idades mais avançadas (Bolígon et al., 2012). Contudo, a inclusão dos efeitos maternos para a análise de dados desta população contribui para a melhor divisão da variância dentro e entre vacas, bem como a partição da variância total, o que resulta em estimativas mais acuradas dos valores genéticos. Além disso, a não inclusão dos efeitos maternos nessa população poderia superestimar a variância genética aditiva direta e, conseqüentemente, os valores genéticos dos animais resultando em menor ganho genético do que o esperado.

As médias de herdabilidade direta apresentaram uma grande queda no início da curva retomando o crescimento logo após. A oscilação permanece até próximo aos 180 dias, apresentando valores médios entre 0,30 e 0,40, e voltando a crescer no final do intervalo até 0,42 (Figura 3.3). Para pesos próximos ao nascimento, foram encontradas herdabilidades superestimadas, assim como descrito por Passafaro et al. (2014), que encontraram herdabilidades superestimadas (acima de 0,60). Estes resultados foram diferentes do relatado por Sousa Junior et al. (2010), em que, a herdabilidade direta foi crescente de próximo a zero até se estabilizar em torno de 0,30 a partir dos 145 dias de idade.

A superestimação da herdabilidade direta no início da trajetória é decorrente dos maiores valores de variâncias direta, como já apresentada na Figura 3.2. No início e no final do intervalo de idades os valores de herdabilidade devem ser avaliados com cautela, pois o ajuste da curva para o efeito genético aditivo direto pode não ter sido adequado, geralmente, pela quantidade de dados coletados não serem suficientes (Passafaro et al., 2014).

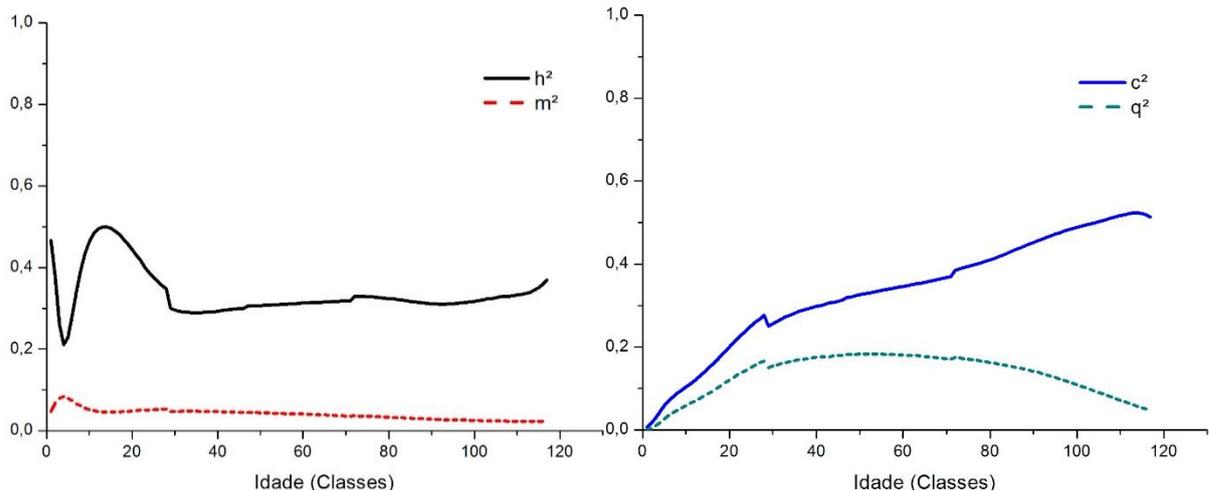


Figura 3.3. Estimativas de herdabilidade direta (h^2) e materna (m^2) e componentes de variância para ambiente permanente de animal (c^2) e materno (q^2) como proporção da variância fenotípica, nas 117 classes de idade, utilizando modelo de regressão aleatória.

Ainda com base nos valores médios de herdabilidade direta, os resultados apontaram valores inferiores ao valor mínimo de 0,58 estimado por Scalez et al. (2014) para o peso pós-desmama de bovinos Nelore em provas de ganho em peso. Uma possível explicação para as diferenças nas herdabilidades com relação ao trabalho de Scalez et al. (2014) pode ser a não inclusão dos efeitos maternos no estudo citado, o que pode ter superestimado a variância aditiva direta.

A variância de ambiente permanente direto, expressa como proporção da variância fenotípica, teve um crescimento acentuado ao longo de todo o intervalo (Figura 3.3). Em geral, a magnitude desse componente foi de média a alta, com valor mínimo de 0,004 e máximo de 0,56. Essa tendência pode ser justificada em razão do caráter cumulativo deste efeito (Schaeffer, 2008). O comportamento não constante para o efeito de ambiente permanente foi observado por Albuquerque & Meyer (2001) e por Nephawe (2004).

Quanto aos valores de herdabilidade materna, estes foram de baixa a média magnitude, com valor mínimo de 0,02 e máximo de 0,07. De modo geral, os valores de herdabilidade materna sugerem que a seleção para este componente implicará em menor resposta do que a

seleção para o efeito direto. Tendências diferentes para esse efeito foram obtidas por Albuquerque & Meyer (2001), Nobre et al. (2003b) e Boligon et al. (2010), com aumento do efeito materno do nascimento aos 205 dias e redução após a desmama.

As médias da variância do efeito de ambiente permanente materno, expressa como proporção da variância fenotípica, apresentaram tendência de crescimento do nascimento até os 300 dias, reduzindo gradativamente deste ponto até o final do intervalo (Figura 3.3).

Em geral, os efeitos maternos neste estudo apresentaram redução após o período da desmama, indicando que tais efeitos têm pouca importância em idades mais avançadas.

Conclusões

Uma função polinomial de Legendre com as ordens seis, três, três e três para os efeitos genéticos direto, efeito materno direto, ambiente permanente do animal e ambiente permanente da mãe, respectivamente, podem ser utilizada para descrever as mudanças nas (co)variâncias ao longo da curva de crescimento dos bovinos Nelore Mocho da região Norte do Brasil.

Referências Bibliográficas

ASSOCIAÇÃO BRASILEIRA DOS CRIADORES DE ZEBU – ABCZ. Estatísticas. [2012] Disponível em: <www.abcz.org.br/conteudo/tecnica/estatisticas.html> Acesso em 15/02/2012.

ALBUQUERQUE, L.G & FARO L. E. Comparação entre os valores genéticos para características de crescimento de bovinos da raça Nelore preditos com modelos de dimensão finita e infinita. **Revista Brasileira de Zootecnia**. v.37, n.2, p.238-246, 2008

ALBUQUERQUE, L.G. & MEYER, K. Estimates of covariance functions for growth from birth to 630 days of age in Nelore cattle. **Journal of Animal Science**, v.79, p.2776-2789, 2001.

BOLIGON, A.A.; ALBUQUERQUE, L.G.; MERCADANTE, M.E.Z.; LÔBO, R.B. Study of relations among age at first calving, average weight gains and weights from weaning to maturity in Nelore cattle. **Revista Brasileira de Zootecnia**, v.39, p.746-751, 2010.

BOLIGON, A. A. et al. Random regression analyses using B-spline functions to model growth of Nelore cattle. **Animal**, v. 6, n. 02, p. 212-220, 2012.

FISCHER, T.M.; VAN DER WERF, J.H.J.; BANKS, R.G. et al. Description of lamb growth using random regression on field data. **Livestock Production Science**, v.89, p.175-185, 2004

GONÇALVES, F.M; PIRES, A.V.; PEREIRA, I.G.; GARCIA, D.A.; FARAH, M.M.; MEIRA, C.T.; CRUZ, V.A.R. Avaliação genética para peso corporal em um rebanho Nelore. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.63, n.1, p.158-164, 2011.

JAMROZIK, J.; SCHAEFFER, L.R. Estimates of genetic parameters for a test day model with random regression for yield traits of first lactation holsteins. **Journal of Dairy Science**, v.80, p.762-770, 1997.

KIRKPATRICK, M.; HECKMAN, N. A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. **Journal of Mathematical Biology**, v.27, p.429-450, 1989.

LAUREANO, M.M.M.; BOLIGON, A.A.; COSTA, R.B.; FORNI, S.; SEVERO, J.L.P.; ALBUQUERQUE, L.G. Estimativas de herdabilidade e tendências genéticas para características de crescimento e reprodutivas em bovinos da raça Nelore. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.63, n.1, p.143-152, 2011.

LÔBO, R.N.B.; BITTNECOURT, T.C.B.S.C.; PINTO, L.F.B. Progresso científico em melhoramento animal no Brasil na primeira década do século XXI. **Revista Brasileira de Zootecnia**, v.39, p. 223-235, 2010.

MEYER, K.; HILL, W.G. Estimation of genetic and phenotypic covariance functions for longitudinal 'repeated' records by restrict maximum likelihood. **Livestock Production Science**, v.47, n.3, p.85-200, 1997

MEYER, K. Scope for a random regression model in genetic evaluation of beef cattle for growth. **Livestock Production Science**, v.86, p.69-83, 2004.

MEYER, K. Estimates of covariance functions for growth of Angus cattle from random regression analyses fitting B-spline functions. **Proceedings of the Association for Advancement of Animal Breeding Genetics**, v.16, p.52-55, 2005.

MEYER, K. “WOMBAT” – Digging deep for quantitative genetic analyses by restricted maximum likelihood. In: World Congress on genetic applied to Livestock Production, 8, 2006, Belo Horizonte. **Proceedings...** Belo Horizonte, 2006, CD-ROM.

MEYER, K. WOMBAT – A program for mixed model Analyses by Restricted Maximum Likelihood. **Journal of Zhejiang University Science B**, v.8, n.11, p.815–821, 2007

NEPHAWE K.A. Application of random regression models to the genetic evaluation of cow weight in Bonsmara cattle of South Africa. *S. Afr. J. Animal Science.*, v.34, p.166-173, 2004.

NOBRE P. R. C.; MISZTAL I.; TSURUTA S.; BERTRAND J. K.; L. SILVA O. C.; LOPES P. S.. Analyses of growth curves of Nelore cattle by multiple-trait and random regression models. **Journal Animal Science**, v.81, p.918-926, 2003.

PASSAFARO, T. L.; FRAGOMENI, B. D. O.; GONÇALVES, D. R.; MORAES, M. M. D.; TORAL, F. L. B. Genetic analysis of body weight in a Nelore cattle herd. **Pesquisa Agropecuária Brasileira**, v. 51, n. 2, p. 149-158, 2016.

SARMENTO, J.L.R.; TORRES, R.A.; PEREIRA, C.S. et al. Avaliação genética de características de crescimento de ovinos Santa Inês utilizando modelos de regressão aleatória. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.58, n.1, p.68-77, 2006.

SCALEZ, D.C.B.; FRAGOMENI, B.O.; PASSAFARO, T.L.; PEREIRA, I.G.; TORAL, F.L.B. Polynomials to model the growth of young bulls in performance tests. **Animal**, v.8, p.370-378, 2014.

SCHAEFFER, L.R. & JAMROZIK, J. Random regression models: a longitudinal perspective. **Journal of Animal Breeding and Genetics**, v.125, n., p.145-146, 2008.

SOUSA JÚNIOR, S. C.; OLIVEIRA, S. M. P.; ALBUQUERQUE, L. G.; BOLIGON, A. A.; MARTINS FILHO, R. Estimação de funções de covariância para características de crescimento da raça Tabapuã utilizando modelos de regressão aleatória. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 39, n. 5, p. 1037-1045, 2010

7 CAPÍTULO IV

Funções de covariâncias sob polinômios B-spline para modelagem do crescimento de bovinos da raça Nelore Mocho da região Norte do Brasil

Resumo: Objetivou-se comparar diferentes modelos de regressão aleatória ajustados por polinômios segmentados do tipo B-spline e ortogonais de Legendre com a finalidade de determinar o mais adequado para descrever mudanças nos parâmetros de avaliação do crescimento de bovinos Nelore. Utilizaram-se 15.148 registros de pesos de 3.115 de bovinos Nelore Mocho, com idade entre 1 e 660 dias, criados na região Norte do Brasil e nascidos entre os anos de 1995 e 2010. Os modelos incluíram, como aleatórios, os efeitos genéticos aditivo direto e aditivo materno, e de ambiente permanente direto e materno e, como fixos, os efeitos de grupo de contemporâneos, a idade da vaca ao parto (efeito linear e quadrático), o polinômio ortogonal de Legendre da idade do animal (regressão cúbica), como covariáveis. O resíduo foi modelado por meio de sete classes heterogêneas. Os efeitos aleatórios foram modelados por meio de funções B-splines. Foram considerados polinômios linear (L), quadrático (Q) e cúbico (C) para cada segmento individual. Polinômios de mesmo grau foram considerados no modelo para todos os efeitos aleatórios, e até cinco segmentos foram utilizados para os efeitos genético direto. Os melhores modelos empregando funções B-spline de diferentes graus foram comparados com um modelo de regressão aleatória sobre polinômios ortogonais de Legendre com ordens 6333 para os efeitos genéticos aditivo direto e aditivo materno, e de ambiente permanente direto e materno, respectivamente. Embora o modelo que utilizou os polinômios ortogonais de Legendre (leg_6333) tenha sido mais parcimonioso, o modelo ajustando funções B-spline cúbicas (bspC_6555), com quatro nós e três segmentos para os efeitos genético aditivo direto e, três nós e dois segmentos para os efeitos genético materno, ambiente permanente de animal e ambiente permanente materno, foi capaz de descrever de forma adequada, com maior velocidade de convergência, a estrutura de (co)variâncias do conjunto de dados de peso de bovinos Nelore criados na região Norte do Brasil.

Palavras-chave: modelos de regressão aleatória, Legendre, trajetória de crescimento, parâmetros genéticos.

CHAPTER IV

Covariance functions on B-spline polynomials for modeling the growth of Nelore Mocho cattle in the Northern region of Brazil

Abstract: The objective of this study was to compare different random regression models adjusted by Legendre type B-spline and orthogonal segmented polynomials with the purpose of determining the best fit to describe changes in Nelore cattle growth parameters. We used 15,148 weight records of 3,115 Nelore Mocho cattle, aged between 1 and 660 days, from the northern region of Brazil, born between 1995 and 2010. The models included, as random, the direct additive genetic effects and maternal additive, and of direct and maternal permanent environment and, as fixed, the effects of group of contemporaries, the age of the cow at birth (linear and quadratic effect), Legendre orthogonal polynomial of the animal's age (cubic regression) as covariables. The residue was modeled by means of seven heterogeneous classes. The random effects were modeled using B-splines functions. Linear (L), quadratic (Q) and cubic (C) polynomials were considered for each individual segment. Polynomials of the same degree were considered in the model for all random effects, and up to five segments were used for direct and maternal genetic effects. The best models employing B-spline functions of different grades were compared with a random regression model on Legendre orthogonal polynomials with orders 6333 for direct additive and maternal additive and direct maternal and permanent environment effects, respectively. The model that considered three-segment cubic B-spline functions for direct and maternal genetic effects, and two segments for the effects of permanent maternal animal environment, was sufficient to model changes in weight variances with the age of animals. Although the model that used Legendre's orthogonal polynomials (leg_6333) has been more parsimonious, the model fitting cubic B-spline functions (bspC_6555) with four nodes and three segments for direct additive genetic effects and three nodes and two segments for the maternal genetic effects, the permanent environment of the animal and the permanent maternal environment, was able to adequately describe, with greater speed of convergence, the (co) variance structure of the Nelore Mocho bovine weight data set created in the North Brazil.

Keywords: Models of random regression, Legendre, growth trajectory, genetic parameters.

Introdução

Os modelos de regressão aleatória com base em polinômios de Legendre vêm sendo empregados no estudo de parâmetros genéticos para características de crescimento de bovinos de corte, como pesos em diferentes idades (MEYER, 2005; SOUSA JÚNIOR et al., 2010; BOLIGON et al., 2012; FERREIRA et al., 2015).

Contudo, vários autores têm destacado que a utilização destes modelos de regressão aleatória sob funções polinomiais de Legendre, pode apresentar algumas propriedades indesejáveis, como a necessidade de utilização de polinômios de alto grau e problemas de inconsistências nas (co)variâncias ao longo da curva (MEYER, 1999; NOBRE et al., 2003; SAKAGUTI et al., 2003), principalmente nas idades em que o número de medidas de pesos é reduzido, como pesos em idades extremas. Além disso, outra possível desvantagem na utilização de polinômios de Legendre, é a dificuldade de convergência com o aumento da ordem dos polinômios (BOLIGON et al., 2010).

Os modelos de regressão aleatória utilizando funções splines podem ser uma opção para modelar pesos até a idade adulta, uma vez que, essas funções são menos suscetíveis a problemas frequentemente observados com polinômios ortogonais de Legendre. Um particular tipo de funções splines são as chamadas B-splines, que são preferidas para a análise de (co)variância devido às suas propriedades numéricas (RUPPERT et al., 2003).

Trabalhos envolvendo modelos de regressão aleatória sob funções B-splines demonstraram maior eficácia sobre os polinômios de Legendre e os modelos de dimensão finita, e podem ser usados para avaliação genética de características de crescimento em bovinos de corte (BALDI et al., 2010; BOLIGON et al., 2012).

No Brasil, os bovinos de corte criados são, em sua maioria, animais zebuínos mantidos em sistemas de criação extensivo e semi-intensivo. Dentre as raças zebuínas destinadas à pecuária de corte, a raça Nelore é a de maior importância, pois conta com um efetivo correspondente a 90% desta parcela (ABCZ, 2016). Portanto, vale ressaltar, que ainda são escassos estudos com modelos de regressão aleatória sob funções B-splines para estimativas de parâmetros genéticos, com base nos dados de crescimento de bovinos da raça Nelore Mocho.

Esse estudo foi desenvolvido com o objetivo de comparar modelos de regressão aleatória ajustados por polinômios segmentados do tipo B-spline e ortogonais de Legendre com a finalidade de estimar aqueles modelos mais adequados para descrever mudanças nos

parâmetros de avaliação do crescimento de bovinos da raça Nelore Mocho criados na região Norte do Brasil.

Material e método

Descrição dos dados

Os dados são provenientes de rebanhos participantes do Programa de Desenvolvimento Ponderal da Associação Brasileira dos Criadores de Zebu (ABCZ). Estes rebanhos estão localizados nos sete estados da região norte do Brasil (Amazonas, Acre, Amapá, Roraima, Rondônia, Pará e Tocantins).

Inicialmente, o banco de dados contava com 44.106 registros de animais. A edição dos dados consistiu em eliminar: registros de pesos com idade acima de 660 dias; animais sem identificação do pai e/ou da mãe; animais com menos de 3 pesagens; e grupos contemporâneos com menos de 3 animais.

A definição de grupo contemporâneo incluiu: sexo, ano de nascimento, estação do ano ao nascimento, ano da pesagem, mês da pesagem, fazenda e condição de criação. As estações de nascimento foram classificadas em chuvosa (dezembro a maio) e seca (junho a novembro). As condições de criação foram: “mamando” e “desmamado”, com regime alimentar à pasto.

Uma análise de consistência dos dados foi feita com finalidade de excluir registros com possíveis erros de medição (*outliers*), através da constatação de uma variabilidade expressiva dentro de cada grupo contemporâneo. Foram excluídos os registros de peso que apresentavam valores fora do intervalo de ± 3 desvios-padrão em relação à média do grupo.

Após edição dos dados, havia 15.148 observações de peso pertencentes a 3.115 animais nascidos entre os anos de 1995 a 2010 e criados em regime alimentar à pasto, com peso médio ao nascer de 28 kg.

O peso médio e o número de registros por idade está representado na Figura 4.1.

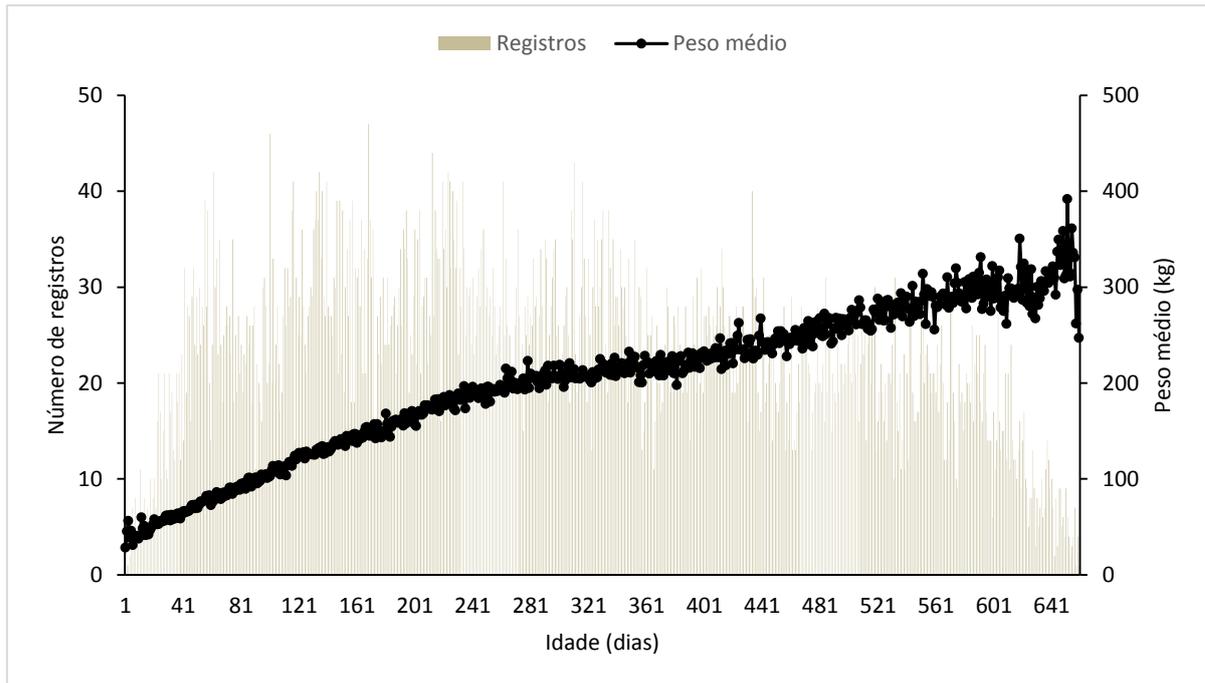


Figura 4. 1. Número de registros e peso médio por idade de bovinos da raça Nelore Mocho criados na região Norte do Brasil.

Os pesos dos animais foram agrupados em 117 classes de idades. Cada classe foi composta a cada intervalo de 5 dias de idade, exceto no início e no final da curva, em que o intervalo de classes foi de 15 dias de idade, devido a diminuição na frequência dos registros nas idades localizadas nos extremos da curva, como se pode observar na Figura 4.1.

A estrutura de dados com número de observações por animal e informações referentes ao pedigree está representada na tabela 4.1.

Tabela 4.1-Resumo da estrutura dos dados de bovinos Nelore Mocho provenientes da região Norte do Brasil.

Informação	Quantidade
Dados	15.148
Animais com observação	3.115
Com 3 observações	712
Com 4 observações	698
Com 5 observações	549
Com 6 observações	659
Com 7 a 10 observações	497
Reprodutores	282
Matrizes	764
Animais na matriz de parentesco	5.678
Grupos contemporâneos	1.819

A edição dos dados e as análises de consistência foram feitas com auxílio do software SAS, versão 9.0 (2003).

Descrição da análise

Para todos os modelos foram incluídos os efeitos genéticos direto e materno, de ambiente permanente direto e materno, além do resíduo. Os efeitos aleatórios foram modelados por meio de funções B-spline.

Foram considerados como efeitos fixos o grupo de contemporâneos, além da idade da mãe ao parto (dias) como covariável (efeitos linear e quadrático), e para modelar a tendência da média populacional foi utilizado polinômio ortogonal de Legendre de quarta ordem.

O resíduo foi considerado em classes heterogêneas, contendo sete classes de variâncias, sendo que as idades foram agrupadas da seguinte forma 1, 2-22, 23-28, 29-46, 47-71, 72-107, 108-117 classes de idade.

O modelo geral pode ser representado por:

$$\mathbf{Y} = \mathbf{Xb} + \mathbf{Z1a} + \mathbf{Z2m} + \mathbf{Z3c} + \mathbf{Z4q} + \mathbf{e}$$

Onde \mathbf{Y} é o vetor das observações, \mathbf{b} é o vetor dos efeitos fixos, \mathbf{a} é o vetor dos coeficientes aleatórios para o efeito genético aditivo direto, \mathbf{m} é o vetor dos coeficientes aleatórios para o efeito genético materno, \mathbf{c} é o vetor dos coeficientes aleatórios para o efeito de ambiente permanente de animal, \mathbf{q} é o vetor dos coeficientes aleatórios para o efeito de ambiente permanente materno, \mathbf{e} é o vetor dos efeitos residuais, \mathbf{X} , $\mathbf{Z1}$, $\mathbf{Z2}$, $\mathbf{Z3}$ e $\mathbf{Z4}$ são as matrizes de incidência. Na forma matricial, o modelo anterior, com suas respectivas pressuposições, pode ser escrito como:

$$v \begin{bmatrix} a \\ m \\ c \\ q \\ \varepsilon \end{bmatrix} = \begin{bmatrix} K_a \otimes A & 0 & 0 & 0 & 0 \\ 0 & K_m \otimes A & 0 & 0 & 0 \\ 0 & 0 & K_c \otimes I & 0 & 0 \\ 0 & 0 & 0 & K_q \otimes I & 0 \\ 0 & 0 & 0 & 0 & R \end{bmatrix}$$

Em que, \mathbf{Y} é um vetor de N observações referentes a Nd animais; \mathbf{b} é um vetor que contém os coeficientes da regressão fixa; \mathbf{a} é um vetor $Ka \times N_D$ de coeficientes de regressão aleatória genéticos aditivos diretos, em que $N_D > Nd$ denota o número total de animais na análise, ou

seja, na matriz de parentesco (5.678); \mathbf{m} é um vetor $K_m \times N_D$ de coeficientes de regressão aleatória genéticos aditivos maternos; \mathbf{c} é um vetor $K_c \times N_d$ de coeficientes de regressão aleatória ambiente permanente de animal; \mathbf{q} é um vetor $K_q \times N_m$ de coeficientes de regressão aleatória de ambiente permanente materno, sendo N_m igual ao número de fêmeas que têm progênies com dados; $\boldsymbol{\varepsilon}$ é um vetor de erros aleatórios; \mathbf{X} , \mathbf{Z}_1 , \mathbf{Z}_2 , \mathbf{Z}_3 e \mathbf{W}_4 referem-se às matrizes de incidência dos efeitos fixos, genético aditivo direto, genético aditivo materno, ambiente permanente de animal, e ambiente permanente materno, respectivamente.

K_a , K_c , K_m e K_q são as matrizes de (co)variâncias entre os coeficientes de regressão aleatórios genético aditivo direto, ambiente permanente de animal, genético aditivo materno e ambiente permanente materno, respectivamente; \mathbf{A} é a matriz dos numeradores do coeficiente de parentesco entre os indivíduos; \mathbf{I}_{N_d} é uma matriz identidade de dimensão N_d ; \mathbf{I}_{N_m} é uma matriz identidade de dimensão N_m ; \otimes é o produto de Kronecker; \mathbf{R} é uma matriz diagonal de variâncias residuais, e sua dimensão depende da estrutura empregada na análise;

A covariância entre os efeitos genéticos aditivos direto e materno foi assumida como sendo igual a zero. As (co)variâncias entre os coeficientes de regressão aleatória genéticos aditivos direto e materno e ambiente permanente de animal e materno, conforme o modelo ajustado, foram estimadas pelo método da máxima verossimilhança restrita (REML), usando o software WOMBAT (Meyer, 2006).

A função B-spline pode ser definida recursivamente. Funções bases de grau $p=0$ tendo valores um para todos os pontos em um dado intervalo, e zero caso contrário.

Para o k -ésimo intervalo dado por nós T_k e $T_k < T_{k+1}$,

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

Funções bases de alto grau, $B_{k,p}$ para $p > 0$, serão então determinadas por funções bases de baixo grau e uma gama de intervalos adjacentes entre os nós. A função pode ser representada:

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

Foram considerados polinômios linear (L), quadrático (Q) e cúbico (C) para cada segmento individual, com funções bases, respectivamente, de graus $p=1, 2$ e 3 .

Polinômios de mesmo grau foram considerados no modelo para todos os efeitos aleatórios. Os m nós foram escolhidos para dividir as classes de idade à pesagem em $m-1$ intervalos equidistantes.

O número de coeficientes de regressão aleatória para modelar a trajetória das funções bases lineares, quadrática e cúbica, foram dados, respectivamente, por m , $m+1$ e $m+2$. A citação dos modelos de regressão aleatória segue o padrão: “BSPX ka kap”, em que X corresponde aos graus dos segmentos polinomiais (L, Q ou C), ka e kap especificam o número de coeficientes de regressão aleatória para os efeitos genético aditivo direto e materno (ka) e de ambiente permanente direto e materno (Kap), respectivamente. Até cinco segmentos foram utilizados para os efeitos genético direto e materno.

Portanto, um modelo tendo a seguinte representação bspQ4433 constará de uma função B-spline quadrática, o número de coeficientes de regressão aleatória igual a 4, tanto para o efeito aleatório genético aditivo como o efeito genético materno, com 3 nós ($m+1$) e 2 segmentos e, 3 coeficientes de regressão aleatória tanto para o efeito ambiente permanente direto e ambiente permanente materno, com 2 nós ($m+1$) e 1 segmento.

Esses modelos empregando funções B-spline foram comparados com um modelo de regressão aleatória sobre polinômios ortogonais de Legendre, de quinta ordem para efeito genético direto e segunda ordem para os efeitos genético materno e de ambiente permanente direto e materno, respectivamente (Leg6333), contendo 7 classes de variâncias residuais.

Os componentes de (co)variâncias e os parâmetros genéticos dos modelos foram estimados pelo método da Máxima Verossimilhança Restrita (REML), utilizando o pacote estatístico Wombat (MEYER, 2006).

A comparação entre os modelos de regressão aleatória foi efetuada pelos critérios de informação de Akaike (AIC) e Bayesiano de Schwarz (BIC) (WOLFINGER, 1993) e pelo exame das variâncias e correlações estimadas para os pesos nas diferentes idades no conjunto de dados.

Os testes de AIC e BIC permitem uma comparação entre os modelos não aninhados e penalizam aqueles com maior número de parâmetros, sendo que o BIC atribui uma penalidade mais rigorosa.

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 o número total de observações, r é o posto da matriz de incidência dos efeitos fixos no modelo e $\log L$ o logaritmo da função de máxima verossimilhança restrita. Menores valores para AIC e BIC indicam melhor ajuste.

Resultados e discussão

Foram apresentados os modelos B-spline, com segmentos lineares, quadráticos e cúbicos, considerando polinômios com até 5 segmentos para os efeitos genético aditivo direto e materno, e comparados com o modelo ajustado por polinômio de Legendre (leg_6333), (Tabela 4.2).

Tabela 4. 2. - Modelos, número de parâmetros (n) do modelo, posição dos nós (classe de idade) para os efeitos genéticos direto e materno, números de segmentos do efeito genético direto (s), critérios de informação de Akaike (AIC) e Bayesiano de Schwarz BIC), para os modelos de regressão aleatória.

Modelos ¹	n	Posição dos nós		s	AIC	BIC
		Genético direto	Genético materno			
bspL_6633	61	1, 24, 47, 70, 93, 117	1, 24, 47, 70, 93, 117	5	87245.3	87702.6
bspL_6655	79	1, 24, 47, 70, 93, 117	1, 24, 47, 70, 93, 117	5	87224.3	87816.5
bspL_6666	91	1, 24, 47, 70, 93, 117	1, 24, 47, 70, 93, 117	5	87144.4	87826.6
bspQ_4444	47	1, 60, 117	1, 60, 117	2	87440.5	87792.9
bspQ_5544	57	1, 39, 77, 117	1, 39, 77, 117	3	87254.0	87681.3
bspQ_5555	67	1, 39, 77, 117	1, 39, 77, 117	3	87221.4	87723.6
bspQ_6555	73	1, 30, 59, 88, 117	1, 39, 77, 117	4	87167.3	87714.6
bspQ_6655	79	1, 30, 59, 88, 117	1, 30, 59, 88, 117	4	87177.8	87770.0
bspQ_6666	91	1, 30, 59, 88, 117	1, 30, 59, 88, 117	4	87150.3	87832.5
bspQ_7777	115	1, 24, 47, 70, 93, 117	1, 24, 47, 70, 93, 117	5	87124.0	87986.2
bspC_6555	73	1, 39, 77, 117	1, 60, 117	3	87145.9	87693.2
bspC_6655	79	1, 39, 77, 117	1, 39, 77, 117	3	87157.2	87749.5
bspC_6666	91	1, 39, 77, 117	1, 39, 77, 117	3	87173.8	87856.0
bspC_7777	115	1, 30, 59, 88, 117	1, 30, 59, 88, 117	4	87140.8	88002.9
bspC_8555	85	1, 24, 47, 70, 93, 117	1, 60, 117	5	87135.2	87772.4
bspC_8855	103	1, 24, 47, 70, 93, 117	1, 24, 47, 70, 93, 117	5	87136.6	87908.8
bspC_8888	139	1, 24, 47, 70, 93, 117	1, 24, 47, 70, 93, 117	5	87056.0	88098.0
Leg_6333	46	-	-	1	87131.2	87476.0

¹ Número de coeficientes de regressão aleatória para os efeitos genético aditivo direto, genético materno, de ambiente permanente de animal e ambiente permanente materno, respectivamente.

Observou-se na Tabela 4.2 que à medida em que o número de segmentos para efeito genético direto foi maior, os valores do critério de informação de Akaike melhoraram, sendo que, a melhora foi maior à medida que se aumentou o grau da função B-spline do modelo.

Estes efeitos foram semelhantes aos relatados por MEYER (2005) e BALDI et al. (2010) utilizando funções B-spline (linear, quadrática e cúbica) para modelar a trajetória de

crescimento do nascimento aos 821 dias de idade em bovinos da raça Angus e do nascimento aos 8 anos de idade de animais da raça Canchim, respectivamente.

De modo geral, os valores do critério de informação de Akaike melhoraram com o aumento no número de parâmetros dos modelos B-spline, evidenciando um melhor ajuste aos dados à medida que o modelo fica mais parametrizado. No entanto, a parametrização excessiva dos modelos aumenta a demanda computacional (memória e capacidade de processamento) e a susceptibilidade a erros numéricos (LEGARRA et al., 2004). Com base nisso, deve-se dar preferência aos modelos mais parcimoniosos.

Neste aspecto, o critério de informação Bayesiano (BIC) pode auxiliar na tomada de decisão, pois este critério penaliza os modelos com um número excessivo de parâmetros. Neste sentido, busca-se a utilização de modelos menos parametrizados, sem que haja uma perda considerável da qualidade do ajuste dos efeitos considerados no estudo, ou seja, quanto menor o valor de BIC, mais parcimonioso é o modelo.

O modelo de regressão aleatória ajustado por polinômios de Legendre (leg_6333) foi o que apresentou o melhor valor de BIC dentre os modelos avaliados, pois, além de possuir um bom ajuste segundo o critério AIC, é o segundo modelo com menor número de parâmetros, atrás, somente, do modelo bspQ_4444.

Os melhores modelos ajustados por funções B-spline, segundo o critério de BIC foram bspL_6633, bspQ_5544 e bspC_6555, para polinômios de graus linear, quadrático e cúbico, respectivamente. As estimativas das variâncias dos efeitos aleatórios para cada modelo, do nascimento aos 660 dias de idade, estão apresentadas na Figura 4.2.

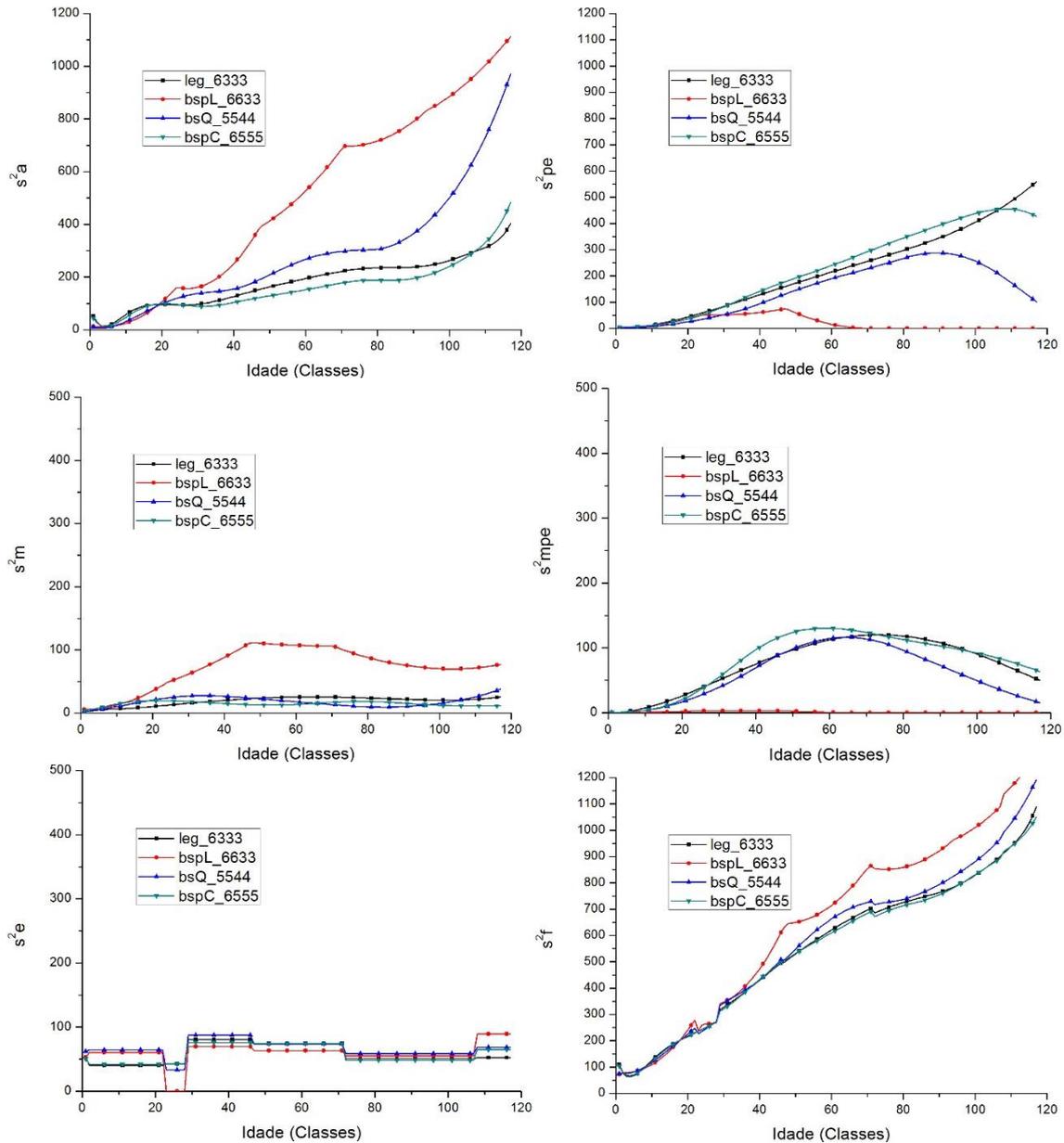


Figura 4. 2. Estimativas de componentes de variância aditiva genética direta (s^2_a), genética materna (s^2_m), ambiente permanente direto (s^2_{pe}), ambiente permanente materno (s^2_{mpe}), residual (s^2_e) e fenotípica (s^2_f), para os modelos Leg6333, bspL_6633, bspQ_5544, e bspC_6555

As estimativas da variância genética aditiva direta apresentaram tendência de crescimento da 1^a a 117^a classe de idade, com grandes diferenças entre a função linear (bspL_6633) e as demais funções. Apesar dos efeitos genéticos serem modelados por 5 segmentos, a função linear não foi capaz de modelar adequadamente as mudanças das (co)variâncias de peso com a idade, resultando em um ajuste grosso e desarmonizado.

Em contraste ao modelo linear, possivelmente por serem funções mais maleáveis, o modelo B-spline cúbico (bspC_6555) descreveu uma curva mais suave e o modelo quadrático

(bspQ_5544) apresentou ligeiras oscilações ao longo da curva, com um crescimento acentuado, resultando em um comportamento inesperado ao final da curva.

Além do efeito genético direto, o modelo B-spline linear superestimou a variância no efeito genético materno, e subestimou a variância nos efeitos de ambiente permanente animal e materno.

A utilização da função B-spline cúbica com 3 segmentos (bspC_6555), bem como a função de Legendre (leg_6333), permitiram melhor controle do aumento da variância genética aditiva direta em relação aos outros modelos. Ambos os modelos estimaram variâncias genéticas aditivas diretas condizentes. Em geral, esses modelos apresentaram estimativas menores em todos os efeitos aleatórios, exceto para o efeito de ambiente permanente materno.

As estimativas de variância genética e de ambiente permanente materno obtidas com os modelos bspC_6555, bspQ_5544 e leg_6333 mostraram tendência semelhante (Figura 4.2). Para esses modelos, maiores variâncias foram estimadas para pesos próximos aos 240 dias de idade (47ª classe de idade), mostrando redução a partir dessa idade e tornando-se nulas em idades avançadas. Resultado semelhante ao obtido por ALBUQUERQUE & MEYER (2001) estudando o comportamento do crescimento de 9.966 animais da raça Nelore, do nascimento aos 630 dias de idade, entre os anos 1981 a 1999.

Dos efeitos aleatórios avaliados, o que apresentou maiores semelhanças entre os diferentes modelos foram estimativa da variâncias residuais, as quais, foram modeladas por 6 classes de variância residual além da classe de variância para peso ao nascer, totalizando 7 classes de variâncias heterogêneas.

As estimativas de variância fenotípica obtidas a partir de modelos de regressão aleatória sobre funções, foram semelhantes e seguiram a mesma tendência crescente entre os modelos.

As estimativas de herdabilidade para o efeito genético aditivo direto obtidas com modelos de regressão aleatória sob funções B-spline e polinômio de Legendre mostraram comportamento semelhante no início da curva. Entretanto, as herdabilidades para o efeito genético aditivo direto obtidas pelos modelos bspQ_5544 e bspL_6633 apresentaram resultados implausíveis (Figura 4.3).

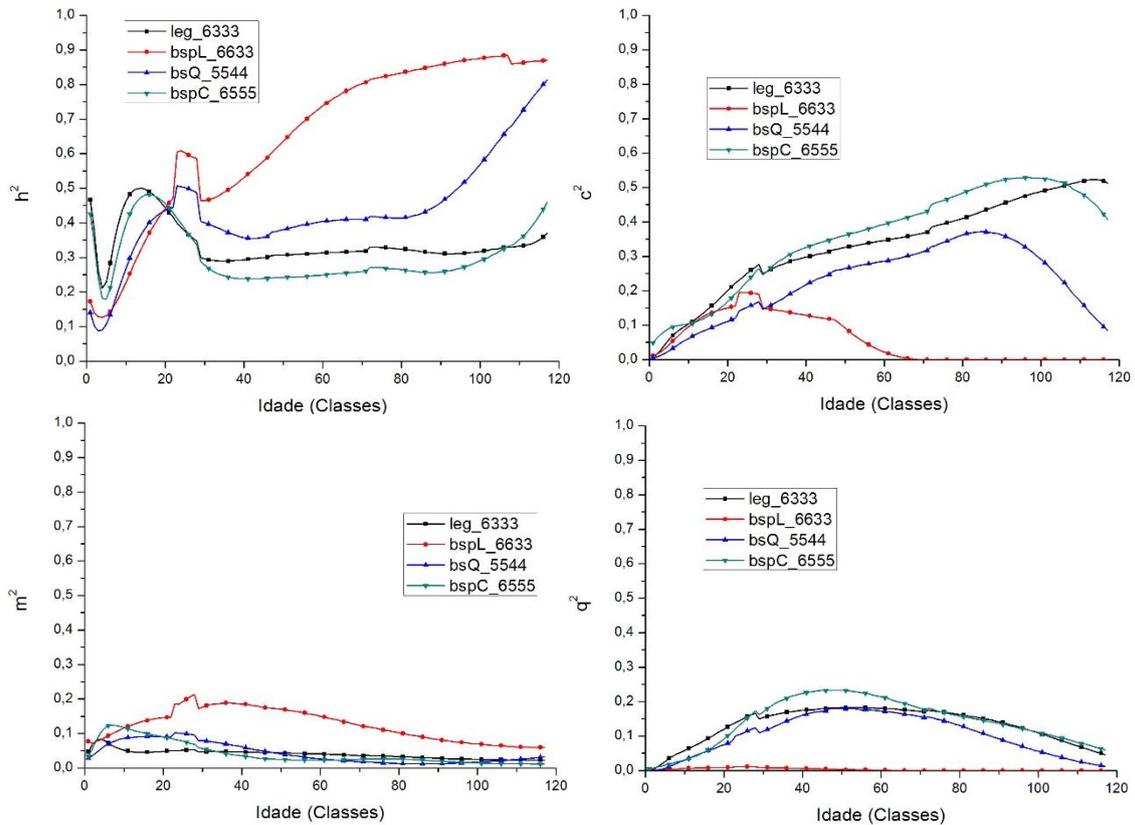


Figura 4. 3. Estimativas de herdabilidade direta (h^2), herdabilidade materna (m^2), estimativa da variância de ambiente permanente de animal (c^2) e materno (q^2) como proporção da variância fenotípica total, obtidas por modelos de regressão aleatória para os modelos de regressão aleatória.

As estimativas da variância de ambiente permanente direto como proporção da variância fenotípica, estimadas pelos modelos bspQ_5544, bspC_6555 e leg_6333, em geral, mostraram a mesma tendência de crescimento (Figura 4.3). O modelo bspL_6633 apresentou comportamento incoerente, com estimativas de variâncias próximas a zero, a partir dos 305 dias de idade (60ª classe). As herdabilidades maternas diminuíram gradualmente até próximo de zero após a 60ª classe de idade, exceto para o modelo bspL_6633.

Para os modelos de regressão aleatória bspQ_5544, bspC_6555 e leg_6333, as estimativas da variância de ambiente permanente materna como proporção da variância fenotípica, seguiram a mesma tendência, com crescimento gradual até próximo à desmama, aproximadamente 240 dias (47ª classe de idade) e um declínio contínuo, posteriormente a desmama, tendendo a zero.

Esse fenômeno é explicado pelas fases do comportamento materno, em que, a contribuição da mãe sobre o fenótipo dos filhos, de acordo com as estimativas da variância de ambiente permanente materna como proporção da variância fenotípica, começou a declinar por

volta da 47ª classe, período da desmama, o que é biologicamente aceitável, pois nos primeiros dias de vida a cria é totalmente dependente dos cuidados maternos (SARMENTO, 2010). Em bovinos de corte, BOLIGON (2012) descreveu comportamento similar, porém o decréscimo do efeito apresentou-se mais tardio. Estimativas com tendências contrárias a estas foram obtidas por SOUSA JUNIOR et al. (2010), em que, a contribuição deste efeito decresceu desde o nascimento com um leve crescimento próximo aos dias 500 dias de idade.

Como se pode observar na Tabela 4.3, para os efeitos genético direto e de ambiente permanente direto, a magnitude das correlações entre os coeficientes de regressão com o intercepto e entre os coeficientes de regressão, diminuiu a medida em que a distância entre os segmentos foi maior. Além disso, a magnitude das variâncias dos coeficientes de regressão diminuiu gradualmente a partir do primeiro coeficiente de regressão até o último coeficiente de regressão, correspondente ao último segmento. Estes resultados estão de acordo com os resultados obtidos utilizando polinômio de Legendre, em que, dentre todos os efeitos aleatórios incluídos no modelo, o intercepto explicou a maior proporção da variância.

Observa-se na Tabela 4.3 que os três primeiros autovalores explicaram a maior parte da variação, portanto, houve autovalores associado à matriz dos coeficientes próximos de zero, o que pode ter sido causado pelo aumento da multicolinearidade devido ao aumento da ordem de ajuste da função empregada, o que sugere redução na dimensionalidade.

Tabela 4. 3. Estimativas de variâncias (diagonal), covariâncias (abaixo da diagonal) e correlações (acima da diagonal) entre os coeficientes de regressão aleatória e autovalores associados à matriz dos coeficientes para o modelo bspC_6555.

	0	1	2	3	4	5	Autovalor
Efeito genético aditivo ($k_a=6$)							Autovalor (%)
0	24516.00	-0.99	0.62	-0.39	0.13	0.03	62.34
1	-5163.20	1101.80	-0.64	0.45	-0.21	0.07	35.73
2	1686.60	-364.80	297.19	-0.10	0.54	0.08	1.68
3	-1222.90	303.39	-33.24	410.59	-0.24	0.84	0.20
4	569.84	-192.02	253.50	-134.46	752.98	-0.47	0.06
5	504.54	295.45	155.65	2038.80	-1545.90	14319.00	0.00
Efeito genético aditivo materno ($k_m=5$)							
0	671.38	-0.97	-0.17	-0.62	-0.43		90.64
1	-215.41	73.46	0.05	0.49	0.64		5.45
2	-22.82	2.32	25.66	-0.22	-0.24		3.90
3	-109.48	28.48	-7.43	46.57	-0.23		0.00
4	-57.69	28.05	-6.26	-7.92	26.36		0.00
Efeito de ambiente permanente animal ($k_c=5$)							
0	1146.80	-0.99	0.61	0.58	0.45		67.26

1	-303.49	82.35	-0.68	-0.52	-0.36	19.17
2	345.41	-102.55	275.64	0.40	0.49	8.59
3	409.98	-98.58	138.33	439.65	0.52	4.98
4	423.86	-92.21	229.22	305.97	781.55	0.00
Efeito de ambiente permanente materno ($k_G=5$)						
0	79.63	-1.00	1.00	1.00	1.00	99.93
1	-40.02	20.20	-1.00	-1.00	-1.00	0.06
2	117.50	-59.32	174.26	1.00	1.00	0.00
3	92.80	-46.83	137.56	108.60	1.00	0.00
4	83.29	-42.00	123.36	97.40	87.37	0.00

Segundo HAIR et al. (2009), a análise de componentes principais é utilizada como método de redução de conjuntos de dados. No contexto de funções tempo-dependentes, os componentes principais são as autofunções da função de covariância (RAMSAY e SILVERMAN, 2005). Cada autofunção expressa padrões de tendência da forma da curva média da população de uma característica longitudinal (KIRKPATRICK e LOFSVOLD, 1992). Contudo, Legarra et al. (2004) alertam para o fato de que a redução da dimensionalidade devido à eliminação de autovalores próximos de zero não é indicada em todos os casos, uma vez que, ao adotar este critério, pode resultar em modelagem simplista ou inadequada

As estimativas de correlações genéticas direta entre pesos do nascimento aos 660 dias, obtidas para o modelo ajustado por funções B-spline cúbicas (bspC_6555) e por funções de Legendre (leg_6333) foram semelhantes. Na Figura 4.4 foram apresentadas as correlações genéticas diretas entre as classes de pesos avaliadas obtidas pelo modelo bspC_6555.

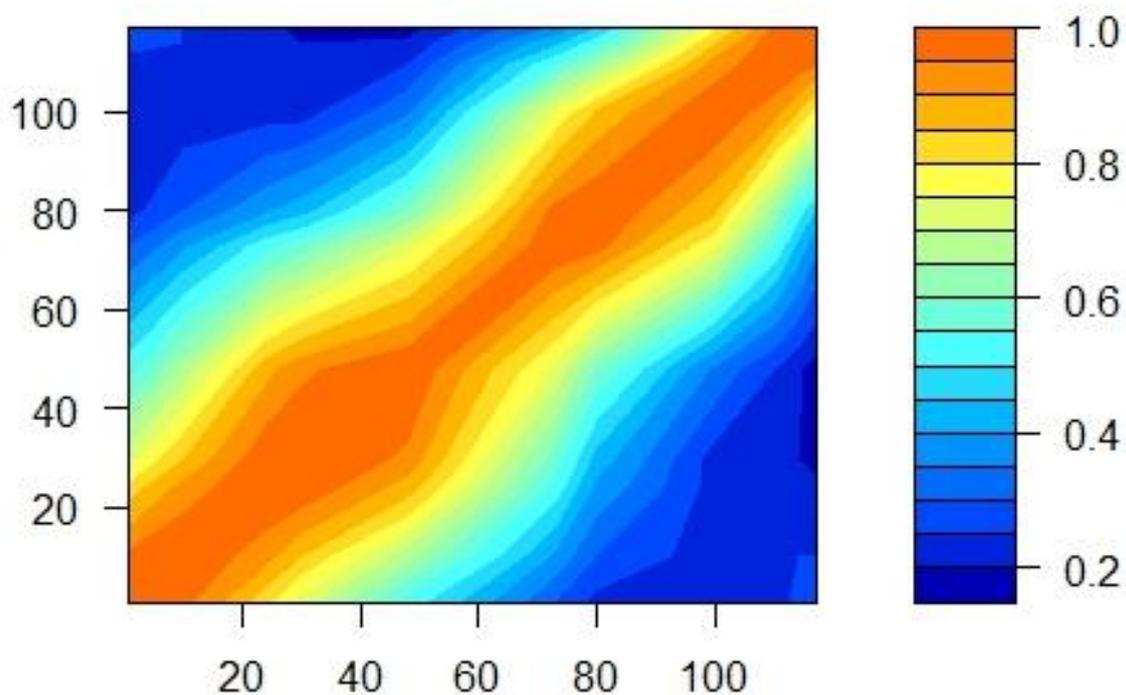


Figura 4. 4. Correlação de Spearman para o valor genético aditivo direto, obtido em diferentes classes de idades, por meio do modelo bspC_6555.

Observa-se na Figura 4.4 que a magnitude da correlação entre os valores genéticos em diferentes idades é representada por um gradiente de cores, como indicado na legenda ao lado do gráfico. As cores quentes, que vão do amarelo ao laranja, representam uma alta correlação e, no sentido contrário, as cores frias, representadas pelas diferentes tonalidades de azul, são indicativas de uma baixa correlação entre os valores genéticos aditivos diretos obtidos em diferentes idades do animal.

Como esperado, a correlação entre os valores genéticos tende a aumentar, quando estes, são preditos em pontos de idades mais próximas entre si. Em contrapartida, a correlação tende a ser nula à medida que a distância entre as idades analisadas tende a ser a máxima (nesse estudo, entre 1 e 660 dias de idade).

Com base nesses resultados pode-se perceber que a correlação entre os valores genéticos diretos no ponto próximo ao peso à desmama (47ª classe) e a idade à maturidade (100ª classe) possui uma correlação de média a baixa magnitude (em torno de 0,4), evidenciando que, a seleção de animais para o peso à maturidade no momento da desmama deve ser feita com cautela.

De modo geral, o modelo de regressão aleatória ajustado por polinômios B-splines cúbicas (bspC_6555), assim como o ajustado pelos polinômios Legendre (leg_6333), apresentaram ajustes semelhantes aos dados de crescimento, sendo que, segundo os critérios AIC e BIC, o modelo leg_6333 proporcionou um ajuste mais parcimonioso.

Segundo BALDI et. al. (2010) para idades em que normalmente são estimados parâmetros genéticos para características de crescimento, como pesos anteriores de seis anos de idade, não houve diferenças importantes nas estimativas obtidas a partir de modelos de regressão aleatória sobre polinômios de Legendre ou funções “B-spline”.

Embora os polinômios B-spline cúbicos tenham necessitado maior número de parâmetros no modelo para estimação dos componentes de (co)variância, estes modelos apresentaram maior velocidade de convergência que os modelos de regressão aleatória sob funções polinomiais de Legendre. Corroborando com esta afirmação, BOLIGON et al. (2012) observaram que, quando se utiliza as funções B-spline, os requisitos computacionais são

menores e a convergência é alcançada de forma mais rápida, devido as equações dos modelos mistos serem esparsas.

Conclusão

O modelo de regressão aleatória ajustado por polinômios ortogonais de Legendre (leg_6333), de ordens seis, três, três, três para os efeitos os efeitos genéticos aditivo direto e materno, ambiente permanente de animal e ambiente permanente materno, respectivamente, foi considerado mais parcimonioso para modelagem de crescimento dos animais estudados.

Um modelo ajustando funções B-spline cúbicas, com quatro nós e três segmentos para os efeitos genético aditivo direto e, três nós e dois segmentos para os efeitos genético materno, ambiente permanente de animal e ambiente permanente materno, descreveu de forma adequada a estrutura de (co)variâncias do conjunto de dados de peso de bovinos Nelore criados na região Norte do Brasil.

Referências Bibliográficas

ASSOCIAÇÃO BRASILEIRA DOS CRIADORES DE ZEBU – ABCZ. Estatísticas. [2012] Disponível em: <www.abcz.org.br/conteudo/tecnica/estatisticas.html> Acesso em 15/02/2012.

ALBUQUERQUE, L.G.; MEYER, K. Estimates of covariance functions for growth from birth to 630 days of age in Nelore cattle. **Journal of Animal Science**, v.79, p.2776-2789, 2001.

BALDI, F. S.; ALENCAR M. M.; ALBUQUERQUE, L.G. Random regression analyses using B-splines functions to model growth from birth to adult age in Canchim cattle. **Journal of Animal Breeding and Genetics**, v. 127, n. 6, p. 433-441, 2010.

BOLIGON, A.A.; ALBUQUERQUE, L.G.; MERCADANTE, M.E.Z.; LÔBO, R.B. Study of relations among age at first calving, average weight gains and weights from weaning to maturity in Nelore cattle. **Revista Brasileira de Zootecnia**, v.39, p.746-751, 2010.

BOLIGON, A. A. et al. Random regression analyses using B-spline functions to model growth of Nelore cattle. **Animal**, v. 6, n. 02, p. 212-220, 2012.

FERREIRA, J. L. et al. Estimação de componentes de (co) variâncias para características de crescimento em bovinos Nelore criados no Trópico Úmido do Brasil por meio de regressão aleatória. **Semina: Ciências Agrárias**, v. 36, n. 3, p. 1713-1724, 2015.

HAIR, J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E. *Multivariate data analysis*. 7 ed. Upper Saddle River: Prentice Hall, 2009. 816 p.

KIRKPATRICK, M.; LOFSVOLD, D. Measuring Selection and Constraint in the Evolution of Growth. *Evolution*, v. 46, p. 954-971, 1992.

LEGARRA, A.; MISZTAI, I.; BERTRAND, J. K. Constructing covariance functions for random regression models for growth in Gelbvieh beef cattle. **Journal Animal Science**, v.82, p. 1564-1571, 2004.

MEYER, K. Estimates of genetic and phenotypic covariance functions for post weaning growth and mature weight of beef cows", **Journal of Animal Breeding and Genetics**, v.116, p.181-205, 1999.

MEYER, K. Estimates of covariance functions for growth of Angus cattle from random regression analyses fitting B-spline functions. **Proceedings of the Association for Advancement of Animal Breeding Genetics**, v.16, p.52-55, 2005.

MEYER, K. "WOMBAT" – Digging deep for quantitative genetic analyses by restricted maximum likelihood. In: World Congress on genetic applied to Livestock Production, 8, 2006, Belo Horizonte. **Proceedings...** Belo Horizonte, 2006, CD-ROM.

NOBRE P. R. C.; MISZTAL I.; TSURUTA S.; BERTRAND J. K.; L. SILVA O. C.; LOPES P. S.. Analyses of growth curves of Nelore cattle by multiple-trait and random regression models. **Journal Animal Science**, v.81, p.918-926, 2003.

PASSAFARO, T. L.; FRAGOMENI, B. D. O.; GONÇALVES, D. R.; MORAES, M. M. D.; TORAL, F. L. B. Genetic analysis of body weight in a Nelore cattle herd. **Pesquisa Agropecuária Brasileira**, v. 51, n. 2, p. 149-158, 2016.

RAMSAY, J.; SILERMAN, B. W. Functional data analysis. New York: Springer, 40 2005. 448p.

RUPPERT, David; WAND, Matt P.; CARROLL, Raymond J. **Semiparametric regression**. Cambridge university press, 2003.

SAKAGUTI, E. S., SILVA, M. A., QUAAS, R. L., MARTINS, E. N., LOPES, P. S., SILVA, L. O. C. Avaliação do crescimento de bovinos jovens da raça Tabapuã, por meio de análises de funções de covariâncias. **Revista Brasileira de Zootecnia**, v. 32, n. 4, p. 864-874, 2003.

SARMENTO, J. L. R et al. Modelos de regressão aleatória na avaliação genética do crescimento de ovinos da raça Santa Inês. **Revista Brasileira de Zootecnia= Brazilian Journal of Animal Science**, p. 1723-1732, 2010.

SOUSA JÚNIOR, S. C.; OLIVEIRA, S. M. P.; ALBUQUERQUE, L. G.; BOLIGON, A. A.; MARTINS FILHO, R. Estimação de funções de covariância para características de crescimento da raça Tabapuã utilizando modelos de regressão aleatória. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 39, n. 5, p. 1037-1045, 2010

WOLFINGER, R. Covariance structure selection in general mixed models. **Communications in Statistics**, Ontario, v. 22, n. 4, p. 1079-1106, 1993.

7 CONSIDERAÇÕES FINAIS

O objetivo dos programas de melhoramento de bovinos de corte é alterar as médias das características de interesse econômico de modo que o lucro nos sistemas de produção aumente. Para isso, modelos estatísticos utilizados nas avaliações genéticas devem ser capazes de identificar corretamente os valores genéticos dos animais candidatos à seleção.

Os modelos de regressão aleatória podem ser utilizados para auxiliar na melhoria na qualidade da modelagem de dados de peso, pois permitem melhorar o aproveitamento dos dados, promovendo incremento na acurácia de avaliação. O desafio é encontrar os modelos mais adequados, dentre os diversos existentes, que são capazes de descrever satisfatoriamente os componentes de covariância dos efeitos aleatórios e, dessa forma, particionar corretamente a variação fenotípica total.

Diversos estudos foram realizados nos últimos anos utilizando funções de covariâncias sob polinômios de Legendre para a modelagem do crescimento de bovinos. Apesar da eficácia comprovada de descrever de forma coerente os efeitos fixos e aleatórios dos modelos, alguns autores afirmam que, em situações complexas, polinômios de alta ordem são requisitados para a modelagem e estes apresentam dificuldades em ajustar a trajetória nos extremos do intervalo, em função da distribuição irregular dos dados ou por artefatos do próprio polinômio de Legendre. Uma solução encontrada para contornar ou minimizar este problema foi a utilização de polinômios splines (segmentados).

Todavia, nesta Tese, os polinômios de Legendre conseguiram explicar o fenômeno sem a necessidade de recorrer a altas ordens, o que pode ser atribuído ao fato do intervalo de idade analisado ter sido muito pequeno (660 dias) e não exigiu ordens polinomiais complexas. Trabalhos mostrando a superioridade de polinômios B-spline sobre polinômios “tradicionais”, geralmente envolvem a avaliação da trajetória de crescimento (ou lactação) em grandes intervalos de tempo, normalmente, acima de seis anos. Portanto, essa é uma situação diferente da demonstrada nesta Tese.

Embora o uso do polinômio B-spline cúbico tenha sido satisfatório na modelagem do crescimento em MRA, e com melhor desempenho computacional, estes foram penalizados pelo grande número de parâmetros gerados pelos modelos. Outro fator determinante, foi a utilização do BIC como critério de avaliação mais importante para escolha de modelos mais harmônicos, embora alguns autores concordem que este critério penaliza demasiadamente os modelos parametrizados.

Todavia, a modelagem de crescimento animal por modelos de regressão aleatória não é uma tarefa simples. Tendo em vista as possibilidades da utilização de polinômios segmentados na modelagem de dados longitudinais, torna-se interessante implementar diferentes metodologias buscando melhorar a acurácia das estimativas em modelos de avaliação genética.