

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE AGRONOMIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ZOOTECNIA**

**MODELO DE REGRESSÃO ALEATÓRIA COMO ALTERNATIVA AO  
MODELO DE LACTAÇÃO NA RAÇA HOLANDESA NO BRASIL**

Alessandro Haiduck Padilha  
Engenheiro Agrônomo/UFRGS

Tese apresentada como um dos requisitos à obtenção do Grau de Doutor em  
Zootecnia  
Área de Concentração Produção Animal

Porto Alegre (RS), Brasil  
Março de 2015

CIP - Catalogação na Publicação

Haiduck Padilha, Alessandro  
MODELO DE REGRESSÃO ALEATÓRIA COMO ALTERNATIVA AO  
MODELO DE LACTAÇÃO NA RAÇA HOLANDESA NO BRASIL /  
Alessandro Haiduck Padilha. -- 2015.  
122 f.

Orientador: Jaime Araújo Cobuci.

Tese (Doutorado) -- Universidade Federal do Rio  
Grande do Sul, Faculdade de Agronomia, Programa de  
Pós-Graduação em Zootecnia, Porto Alegre, BR-RS, 2015.

1. Regressão aleatória. 2. Modelo de lactação aos  
305 dias. 3. Valores genéticos. 4. Confiabilidade.  
I. Cobuci, Jaime Araújo, orient. II. Título.

ALESSANDRO HAIDUCK PADILHA  
Engenheiro Agrônomo e  
Mestre em Zootecnia

## TESE

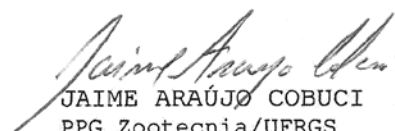
Submetida como parte dos requisitos  
para obtenção do Grau de

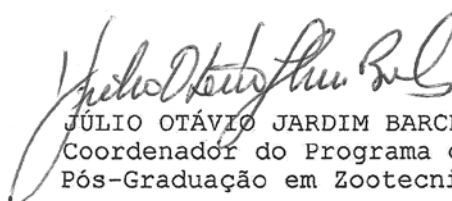
### DOUTOR EM ZOOTECNIA


Programa de Pós-Graduação em Zootecnia  
Faculdade de Agronomia  
Universidade Federal do Rio Grande do Sul  
Porto Alegre (RS), Brasil


Aprovado em: 03.03.2015  
Pela Banca Examinadora

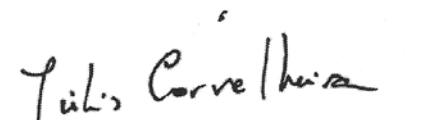
Homologado em: 17.06.2015  
Por


  
JAIME ARAÚJO COBUCI  
PPG Zootecnia/UFRGS  
Orientador

  
JÚLIO OTÁVIO JARDIM BARCELLOS  
Coordenador do Programa de  
Pós-Graduação em Zootecnia

  
JOSÉ BRACCINI NETO  
PPG ZOOTECNIA-UFRGS

  
CLAUDIO NAPOLIS COSTA  
EMBRAPA

  
JULIO GIL VALE CARVALHEIRA  
UNIVERSIDADE DO PORTO

  
PEDRO ALBERTO SELBACH  
Diretor da Faculdade de Agronomia

## **Dedicatória**

**Dedico esse trabalho à esposa Sheila Andréia Carvalho que tem me apoiado nos momentos mais difíceis e ao nosso gato amarelo resgatado da rua, o Juju, cuja companhia tanta alegria tem proporcionado.**

## MODELO DE REGRESSÃO ALEATÓRIA COMO ALTERNATIVA AO MODELO DE LACTAÇÃO NA RAÇA HOLANDESA NO BRASIL<sup>1</sup>

Autor: Alessandro Haiduck Padilha

Orientador: Prof. Jaime Araújo Cobuci

### RESUMO

O presente estudo teve como objetivos comparar: (i) modelos de regressão aleatória ajustados por polinômios de Legendre considerando diferentes de registros de produção no dia do controle por lactação, (ii) qualidade do ajuste dos modelos de regressão aleatória e modelos de lactação aos 305 dias para produção de leite, (iii) qualidade do ajuste dos modelos de regressão aleatória e modelos de lactação aos 305 dias para produção de gordura e proteína. Foram usados dados de produção de vacas de primeira lactação coletados pelos Serviços de Controle Leiteiro e Genealógico da Associação Brasileira de Criadores de Bovinos da Raça Holandesa (ABCBRH) compreendidos entre 1990 e 2011. Foram formadas quatro estruturas de dados ao restringir vacas com pelo menos 4, 6, 8 e 10 e máximo de 11 registros de produção de leite no dia do controle por lactação. Nessas estruturas foram usados modelos de regressão aleatória com polinômios de Legendre de terceira a quinta ordem. Paralelamente, uma base ou estrutura com pelo menos 6 registros de produção de leite, de gordura e proteína por lactação e uma base com registro de produção de leite, gordura e proteína em até 305 dias da lactação foram formados, sendo usados modelos de regressão aleatória de quarta e quinta ordem e modelos de lactação aos 305 dias, respectivamente. Todas as análises foram realizadas por meio da máxima verossimilhança restrita pelo programa REMLF90 nos sistemas IBM, ICE e SGI do CENAPAD-SP. Os valores de AIC, BIC,  $-2\text{LogL}$  e variância residual foram menores para os modelos de quinta ordem, porém os primeiros três ou quatro autovalores explicaram mais de 99% da variação total nos modelos com quarta e quinta ordem, conforme a estrutura de dados. Correlações de Spearman dos valores genéticos para P305 entre modelos com diferentes ordens polinomiais variaram de 0,99 a 1,00. Confiabilidades médias de valores genéticos de touros foram de 0,82, 0,80, 0,80 e 0,64 para estruturas com 4, 6, 8 e 10 registros. O ganho médio na confiabilidade dos valores genéticos para touros foi de 4% a 17% (leite), de 3% a 16% (gordura), e de 6 a 26% (proteína) maiores do que aqueles estimados pelo modelo de lactação. A diferenciação nas estruturas de dados alterou a confiabilidade dos valores genéticos de touros, conforme aumentou a restrição nos dados. Modelos de regressão aleatória são mais acurados que modelos de lactação para ajustar registros de produção de leite, gordura e proteína de animais da raça Holandesa no Brasil.

---

<sup>1</sup>Tese de Doutorado em Zootecnia - Produção Animal, Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil. ( 116 p.) Janeiro, 2015.

## **RANDOM REGRESSION MODEL AS ALTERNATIVE TO THE LACTATION MODEL FOR HOLSTEIN CATTLE IN BRAZIL<sup>1</sup>**

Author: Alessandro Haiduck Padilha

Adviser: Prof. Jaime Araújo Cobuci

### **ABSTRACT**

This study aimed to compare: (i) random regression models adjusted for Legendre polynomials considering different test day records by lactation, (ii) goodness of fit of random regression and lactation models for estimating breeding values for 305day milk yield, (iii) goodness of fit of random regression models and lactation models for estimating breeding values for 305day fat and protein yield. Thus data consisted of milk yield collected by the technicians of the Milk Control and Genealogy of the Brazilian Association of Holstein Breeders (ABCBRH) between 1990 and 2011. Four structures of subsets were formed by restricting cows with at least 4, 6, 8 and 10 test day milk yield records in lactation. For these structures, random regression models with Legendre polynomials of third to fifth orders were used. At same time, a base with 6 test day milk yield records for fat and protein in lactation and a base with records of total lactation yield of milk, fat and protein yields at 305 days. For these bases, random regression models of fourth and fifth order Legendre polynomials and 305 day-lactation models were used, respectively. All analyzes were performed using restricted maximum likelihood by REMLF90 program in IBM, ICE and SGI CENAPAD-SP systems. The values of AIC, BIC, -2LogL and residual variance were lower for fifth order Legendre polynomials but three or four eigenvalues explained over 99% of total variation in models of fifth and fourth orders, according to the structures of data. Spearman correlations of breeding values for Y305 between models with different polynomial orders were 0.99 and 1.00. The average reliability of breeding values of bulls was 0.82, 0.82, 0.80 and 0.64 for structures 4, 6, 8 and 10 test-days, respectively. The average gain in reliability of breeding values of bulls was between 4% and 17% (milk), 3% and 16% (fat), 6 and 26% (protein) higher than that estimated for the 305-day lactation model. The differentiation in the structures of data influenced the reliability of breeding values of bulls, according to the increasing in the restriction of test days. Random regression models are more accurate than lactation models to adjust milk, fat and protein yield records of Holstein cattle in Brazil.

---

<sup>1</sup>Doctoral thesis in Animal Science, Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil. (116 p.) Janeiro, 2015.

## SUMÁRIO

	Página
<b>1. CAPÍTULO I</b> .....	15
1.1 INTRODUÇÃO GERAL .....	17
1.2 REVISÃO BIBLIOGRÁFICA .....	19
1.2.1 Modelos para avaliação genética da raça Holandesa.....	19
1.2.2 Comparação entre modelos.....	24
1.3 HIPÓTESES.....	26
1.4 OBJETIVOS.....	27
<b>2. CAPÍTULO II - Selecting the suitable random regression model for genetic evaluations of Brazilian Holstein cattle ...</b>	29
Abstract .....	31
Keywords .....	32
Implications .....	32
Introduction .....	33
Material and Methods .....	34
Results .....	38
Discussion .....	40
Conclusion .....	44
Acknowledgements .....	44
References .....	44
Figures and Tables .....	49
<b>3. CAPÍTULO III - Random regression models are suitable to substitute the traditional 305-day lactation model in genetic evaluations of Holstein cattle in Brazil.....</b>	55
Abstract .....	57
Keywords .....	58
Introduction .....	58
Material and Methods .....	59
Results .....	63
Discussion .....	65
Conclusion .....	70
Acknowledgements.....	70
References .....	70
Figures and Tables .....	74
<b>4. CAPÍTULO IV - Random regression and 305-day lactation models for fat and protein yields in genetic evaluations of Holstein cattle in Brazil</b> .....	82
Abstract .....	84
Keywords .....	84
Introduction .....	84
Material and Methods .....	86

Results .....	91
Discussion .....	93
Conclusion .....	97
Acknowledgements.....	97
References .....	98
Figures and Tables .....	102
<b>5. CAPÍTULO V – CONSIDERAÇÕES FINAIS .....</b>	<b>108</b>
<b>6. REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>111</b>
<b>7. VITA .....</b>	<b>122</b>



## RELAÇÃO DE TABELAS

### 2. CAPÍTULO II – Selecting the suitable random regression model for genetic evaluations of Brazilian Holstein cattle

Table 1 Description of databases .....	49
Table 2 Values of tests obtained by random regression models in different structures of datasets.....	49
Table 3 Eigenvalues ( $\lambda_i$ ) of the additive genetic (co)variance matrix and the proportion of total variance (%) estimated from Legendre polynomials.....	50
Table 4 Eigenvalues ( $\lambda_i$ ) of the permanent environmental (co)variance matrix and the proportion of total variance (%) estimated from Legendre polynomials.....	50
Table 5 Average, minimum (min) and maximum (max) additive genetic ( $r_g$ ) and permanent environmental ( $r_{pe}$ ) correlations between test days on 6 to 305 days .....	51
Table 6 Spearman rank correlations of breeding values (bulls/cows) for 305-day milk yield between random regression models.....	51
Table 7 Average, standard deviations, minimum and maximum of reliabilities ( $r^2$ ) of estimated breeding values for 305-day yield (Y305) of bulls with more than 10 progenies in different structures .....	52

### 3. CAPÍTULO III – Random regression models are suitable to substitute the traditional 305-day lactation model in genetic evaluations of Holstein cattle in Brazil

Table 1. Number of parameters (p), estimates of the maximum of the likelihood function (-2 Log L), Akaike's information criterion (AIC) and Bayesian information criterion (BIC), change in log likelihood and chi-square statistics ( $\chi^2$ ) for random regression models using Legendre polynomials. ....	75
Table 2. Estimates of additive genetic ( $\sigma^2_a$ ), residual ( $\sigma^2_e$ ) and permanent environmental ( $\sigma^2_{pe}$ ) variance components and heritability coefficients ( $h^2$ ) for 305-day yield trait (305MY) estimated from LM, RRM <sub>4</sub> and RRM <sub>5</sub> models.....	75
Table 3. Standard deviations (kg) of EBVs for 305MY estimated by RRM <sub>4</sub> and RRM <sub>5</sub> models, and percentage of change (between brackets) with respect to standard deviations by LM model for bulls and cows according to progeny size and number of test days.....	76
Table 4. Spearman rank correlation between EBVs for 305MY estimated from LM and EBVs for 305MY from RRM <sub>4</sub> and RRM <sub>5</sub> models for bulls and cows according to progeny size and number of TD.....	76
Table 5. Spearman rank correlation ( $p < 0.0001$ ) between EBVs for 305MY estimated from LM and EBVs for 305MY from RRM <sub>4</sub> and RRM <sub>5</sub> models for bulls with progeny size higher than 49 and cows selected for 305MY.....	77
Table 6. Ranks of EBVs of the ten top bulls with progeny size higher than 49 selected by EBVs for 305MY estimated from RRM <sub>4</sub> and RRM <sub>5</sub> and respective ranks from LM .....	77
Table 7. Classes of reliability ( $r^2$ ) of EBVs and the number of bulls for LM and for RRM <sub>4</sub> and RRM <sub>5</sub> models with percentage of increasing in the number of bulls in the respective class compared to LM (brackets) .....	78

Table 8. Reliabilities and their standard deviations of EBVs for 305MY estimated from LM and the percentage of gain with range in brackets by RRM<sub>4</sub> and RRM<sub>5</sub> models for bulls and cows according progeny size and number of TD ..... 78

#### **4. CAPÍTULO IV – Random regression and 305-day lactation models for fat and protein yields in genetic evaluations of Holstein cattle in Brazil**

Table 1 Number of estimated parameters ( $p$ ), maximum likelihood function ( $2 \text{ Log L}$ ), Akaike's information criterion (AIC), Bayesian information criterion (BIC), likelihood ratio test (LRT) and chi-square statistics ( $\chi^2$ ) for random regression models ..... 104

Table 2 Eigenvalues ( $\lambda_i$ ) of the additive genetic (co)variance matrix and the proportion of total variance (%) estimated from Legendre polynomials..... 104

Table 3 Estimates of heritability for 305-day fat and protein yields estimated from lactation and random regression ..... 105

Table 4 Spearman rank correlations of EBVs for 305-day fat yield between LMF and RRF4 and between LMP and RRP4 of bulls and cows according to progeny size and number of test days ..... 105

Table 5 Average percentage of increasing in reliability (range in parenthesis) of EBVs for 305-day yield estimated from random regression model compared to lactation model ..... 106

## RELAÇÃO DE FIGURAS

### **2. CAPÍTULO II – Selecting the suitable random regression model for genetic evaluations of Brazilian Holstein cattle**

Figure 1 Heritability of test days on days in milk estimated from random regression models with Legendre polynomials from third to sixth order (M3, M4 and M5) applied to structures 4, 6, 8 and 10 TD ..... 53

### **3. CAPÍTULO III – Random regression models are suitable to substitute the traditional 305-day lactation model in genetic evaluations of Holstein cattle in Brazil**

Figure 1 - Heritability, additive genetic and permanent environmental variances of test day milk yields on days in milk estimated from random regression models fitted by fourth (RRM<sub>4</sub>) and fifth (RRM<sub>5</sub>) order Legendre polynomials ..... 79

Figure 2 – Genetic correlation estimates (left) and permanent environmental correlation (right) between test-day milk yield along days in milk..... 79

Figure 3 - Trajectory of test day breeding values on days in milk of five top bulls selected for 305MY breeding values estimated from LM and RRM4 models... 80

### **4. CAPÍTULO IV – Random regression and 305-day lactation models for fat and protein yields in genetic evaluations of Holstein cattle in Brazil**

Figure 1 Heritability, additive genetic (AG) and permanent environmental (PE) variances (x 1,000) of test-day fat (left) and protein (right) yields estimated from random regression models fitted by fourth and fifth-order Legendre polynomials ..... 103

Figure 2 Genetic correlation (left) and permanent environment correlations (right) between test-days on days in milk for fat and protein yields..... 103

## RELAÇÃO DE ABREVIATURAS E SÍMBOLOS

AIC	Critério de informação de Akaike
BIC	Critério de informação Bayesiano
DIM	Dias em lactação
g	Gramas
g <sup>2</sup>	Gramas ao quadrado
h <sup>2</sup>	Herdabilidade
kg	Quilogramas
kg <sup>2</sup>	Quilogramas ao quadrado
log (L)	Função da máxima verossimilhança
LM	Modelo de lactação
LMF	Modelo de lactação para produção de gordura do leite
LMP	Modelo de lactação para produção de proteína do leite
M3	Modelo de regressão aleatória ajustado por polinômios de Legendre de ordem 3
M4	Modelo de regressão aleatória ajustado por polinômios de Legendre de ordem 4
M5	Modelo de regressão aleatória ajustado por polinômios de Legendre de ordem 5
P305	Produção acumulada aos 305 dias
PEV	Variância do erro de predição
PS <sub>1</sub>	Primeira medida para a persistência
PS <sub>2</sub>	Segunda medida para a persistência
PS <sub>3</sub>	Terceira medida para a persistência
PS <sub>4</sub>	Quarta medida para a persistência

PS <sub>5</sub>	Quinta medida para a persistência
PS <sub>6</sub>	Sexta medida para a persistência
PS <sub>7</sub>	Sétima medida para a persistência
PS <sub>8</sub>	Oitava medida para a persistência
PS <sub>9</sub>	Nona medida para a persistência
r <sup>2</sup>	Confiabilidade dos valores genéticos
SEP	Predição do erro padrão
TDM	Test-daymodel ou modelo no dia do controle
TRV	Variância residual total
TDMY	Produção de leite no dia do controle
TDFY	Produção de gordura no dia do controle
TDPY	Produção de proteína no dia do controle
χ <sup>2</sup>	Chi-quadrado



## 1. CAPÍTULO I





## 1.1 INTRODUÇÃO GERAL

A produção leiteira, no Brasil, cresceu mais de 37% na última década (FAO, 2014) em parte devido a melhorias no padrão genético animal. Se, por um lado, a média nacional é de aproximadamente 5 litros por vaca por dia (ANUALPEC, 2011), por outro, há vacas da raça Holandesa cuja produção pode chegar a mais de 40 litros (Aikman et al., 2008), o que significa, comparativamente, sair de um potencial de aproximadamente 2.000 para em torno de 10.000 litros de leite numa única lactação (Dobson et al., 2007). O aumento da eficiência dos sistemas de produção na raça Holandesa depende muito das metodologias utilizadas para avaliação genética dos animais.

A avaliação genética nacional dos animais da raça Holandesa, assim como em outros países, ainda é realizada por um modelo animal que utiliza registro de produção de leite acumulada aos 305 dias, ou seja, pelo modelo de lactação aos 305 dias (Kim et al., 2009; Bilal & Khan, 2009; Biassus et al., 2011; Costa et al., 2012). No entanto, há alguns países que já vem utilizando modelos que utilizam registros de produção de leite no dia do controle, via metodologia da regressão aleatória (Muir et al., 2007). A regressão aleatória, cuja estrutura de covariância entre produções no dia do controle é ajustada por polinômios de Legendre, tem sido avaliada por vários pesquisadores e é atualmente apresentada como a melhor metodologia para avaliação genética de bovinos leiteiros (Strabel et al., 2004; Muiret al., 2007; Dzomba et al., 2010; Khorshidie et al., 2012; Çankaya et al., 2014).

As principais vantagens do modelo que utiliza regressão aleatória quando comparada com o modelo de lactação aos 305 dias são a flexibilidade de levar em consideração os componentes genéticos e aditivos do formato da curva de lactação sem a necessidade de uso de fatores de extensão/projeção para lactações incompletas e a possibilidade da avaliação de características como a persistência da lactação (Strabel et al., 2004; Costa et al., 2008). Apesar dessas vantagens em relação ao modelo de lactação aos 305 dias, ainda não há um consenso na literatura sobre utilização de lactações incompletas ou em cursona análise (ou de número mínimo de controles leiteiros por lactação na análise), sobre a melhor medida de persistência, sobre a utilização de registros de contagem de células somáticas ou sobre a ordem dos polinômios de Legendre mais apropriados para serem adotados na avaliação genética da raça Holandesa no Brasil (Melo et al., 2007; Biassus et al., 2010; Cobuci et al., 2012).

Assim, o foco da pesquisa foi comparar modelos de regressão aleatória na predição de valores genéticos, utilizando diferentes estruturas de bases de dados ao restringir na base original vacas que continham pelo menos 4, 6, 8 e 10 registros de produção de leite por lactação, comparar diferentes

modelos de regressão aleatória ajustados por polinômios de Legendre com o modelo de lactação aos 305 dias atualmente utilizado no Brasil, para características de produção de leite, gordura e proteína.

## 1.2 REVISÃO BIBLIOGRÁFICA

### 1.2.1 Modelos para avaliação genética da raça Holandesa

Os serviços de controle leiteiro no mundo registram como informação básica a produção de leite e componentes (gordura e proteína) produzidos por uma vaca durante 24 horas em um dado dia da lactação, o que é chamado de produção no dia do controle. Segundo Mrode et al. (2013), no Canadá, o governo canadense possuía dois planos de registro: registrar todas as produções diárias, gerando altos custos, ou registrar a cada 30 dias. De qualquer forma, as produções no dia do controle eram padronizadas por meio de fórmulas cumulativamente aos 305 dias, mesmo para vacas com lactações incompletas ou em curso.

No Brasil, o controle leiteiro oficial tem sido realizado mensalmente a intervalos de 15 a 45 dias, sendo que o primeiro controle da lactação é realizado entre o 6º e 75º dia do parto (BRASIL, 1986; Costa et al., 2012). A Associação Brasileira de Criadores de Bovinos da Raça Holandesa (ABCBRH) foi fundada em 1934 e, por meio dos técnicos do Serviço de Classificação Linear, Controle Leiteiro e Genealógico, tem coletado dados para serem usados na avaliação genética da raça Holandesa (Costa, 2005; Ferreira et al., 2002; BRASIL, 1986; Costa et al., 2012). Cada pesagem mensal de leite, com seu respectivo resultado da análise de componentes (gordura ou proteína) ou contagem de células somáticas são denominados registros no dia do controle. Os registros no dia do controle de uma vaca, que normalmente numa lactação completa pode conter aproximadamente 10 registros, são então padronizados num único registro aos 305 dias, que representa toda a produção de leite de uma lactação por meio de fatores de correção/projeção, como os detalhados por Everett & Carter (1968). Esses dados padronizados de produção aos 305 dias (P305) têm sido usados num modelo animal de lactação para avaliação genética oficial de animais Holandês em países como Brasil ou Coréia, por exemplo. Os fatores de correção costumam superestimar a produção de leite acumulada até o primeiro controle, assim como para 305 dias, por isso, alguns autores propuseram a utilização de novos fatores de correção numa tentativa de diminuir o viés (Ferreira et al., 2002; Morais Junior et al., 2002; Çilek, 2008). Além disso, conforme sugerido por Çilek (2008), fatores propostos para uma situação não deveriam ser usados para outros rebanhos, raças ou climas diferentes daqueles para os quais foram propostos, devendo, portanto ser calculados novos fatores para cada caso. Em segundo lugar, cada produção no dia do controle é afetada por diferentes fatores (Bilal e Khan, 2009), tais como raça, região do país, grupo contemporâneo ou de manejo dentro de rebanho (Reents et al., 1995; Jamrozik et al., 1997a), dia do ano, número de lactações

(Jamrozik et al., 1997b; De Roos et al., 2001), idade ao parto (Kaya et al., 2003), mês de parto (Bormann et al., 2003), dias em lactação (Hamed, 1995; Kaya et al., 2003) e número de ordenhas diárias (Wiggans, 1986). O modelo de lactação para avaliar a P305 não leva em consideração as mudanças no ambiente durante a lactação até atingir os 305 dias. Além disso, lactações incompletas, por quaisquer motivos, são projetadas até os 305 dias ou truncadas caso ultrapassem esse período (Bilal e Khan, 2009).

No mundo todo, outras propostas têm sido desenvolvidas para analisar dados de produção no dia do controle, sem pré-ajustamentos, numa nova abordagem metodológica (Ptak e Schaeffer, 1993). Dentre as abordagens relatadas na literatura (Togashi et al., 2004; Sawalha et al., 2005; Hollanda et al., 2011), podemos citar os seguintes:

- a) Análise unicaracterística ou univariada com modelo de repetibilidade: assume variâncias genéticas constantes sobre as idades e correlação genética de uma unidade (1) entre os registros tomados em diferentes idades;
- b) Análise multicaracterística ou multivariada: cada registro de produção em idades diferentes é tratado como características diferentes ou arbitrariamente divide a amplitude de idades em intervalos e trata as medidas de diferentes intervalos como características diferentes.
- c) Ajustam a curva de lactação aos dados e analisam os parâmetros da curva como características diferentes;
- d) Aplicam modelo de repetibilidade autorregressivo: assume uma estrutura de covariância simétrica composta entre as produções no dia do controle;
- e) Aplicam modelos de regressão fixos e/ou aleatórios: assumem que as variâncias entre registros tomados em idades diferentes apresentam correlações diferentes e menores que uma unidade.
- f) Funções spline (polinômios segmentados): A lactação é dividida em secções, assumidas como tendo mudança linear, por localizações denominadas de nós (knots).

Modelos que consideram a produção no dia do controle (ou Test-Day Models) têm sido aceitos como substitutos do modelo tradicional de lactação aos 305 dias (Muiret et al., 2007; Kim et al., 2009; Çankaya et al., 2014). Diferentes PLDC têm sido propostos para avaliar dados longitudinais, conforme revisado por Jensen (2001). Modelos de produção no dia do controle já eram usados desde 1985 na Austrália para criar índices para vacas (Mrode et al., 2013). Em 1990, o modelo animal foi implementado no Canadá devido aos computadores terem ficado menos caros e se tornado mais comum (Mrode et al., 2013). Em 1982, Henderson (1982), propôs o uso de modelos de regressão aleatória para avaliações genéticas de características medidas ao longo do tempo. Em 1994, Schaeffers e Dekkers (1994) introduziram o conceito do modelo de regressão aleatória. A regressão aleatória significa que cada vaca tem sua própria curva de lactação com sua forma própria, a qual é comparada a curva de regressão fixa correspondente a subclasse de efeitos fixos (Mrode, 2005). Jamrozik et al. (1997) e Werf et al. (1998) relataram que modelos de

regressão aleatória foram mais apropriados para estimar parâmetros genéticos no dia do controle para leite comparado aos modelos de repetibilidade. Schaeffer (2004) revisou as aplicações dos modelos de regressão aleatória nas avaliações genéticas. Assim, modelos de regressão aleatória não só têm sido considerados mais promissores e sendo estudados como substitutos aos modelos de lactação aos 305 dias, mas também vem sendo empregados oficialmente em programas de melhoramento de alguns países (Muir et al., 2007; Kim et al., 2009; Dzomba et al., 2010; Bignardi et al., 2011; Cobuci et al., 2011; Yamazaki et al., 2013; Aliloo et al., 2014).

Esses modelos de regressão apresentam muitas vantagens comparativamente aos modelos pré-ajustados aos 305 dias e também aos outros modelos que consideram a produção no dia do controle, como por exemplo:

- a) Considera fatores ambientais que podem afetar diferentemente as vacas durante a lactação, além de permitir considerar curva de lactação diferente para cada vaca (Swalve, 2000; Zavadilová et al., 2005), o que significa numa definição mais precisa dos grupos contemporâneos de animais (Costa et al., 2005);
- b) Permite que vacas sejam avaliadas com base em qualquer número de registros no dia do controle, o que significa avaliações de lactações em curso ou parciais, reduzindo o intervalo de gerações (Swalve, 2000);
- c) Evita a necessidade de pré-ajustamento da curva de lactação para 305 dias e os vieses decorrentes dessa abordagem (Kim et al., 2009; Bilal & Khan, 2009);
- d) Minimiza problemas associados com uma estrutura simplificada dos modelos de repetibilidade ou a superparametrização nos modelos multicaracterística (Jamrozik & Schaeffer, 1997; Meyer, 1998a; Pool et al., 2000; Swalve, 2000).

No entanto, não há consenso sobre a estrutura de covariância entre os registros no dia do controle que melhor modelam as curvas de lactação, assim como o número de registros necessários para uma análise precisa, pois podem diferir conforme as condições de cada programa de avaliação genética de cada país, dependendo da raça e da quantidade de dados de animais registrados disponíveis.

O ajustamento da trajetória ao longo do tempo pode ser realizado por meio de diferentes funções contínuas (Pereira et al., 2010; Bignardi et al., 2009). As funções que se destacam na literatura apresentam vários tipos de abordagens, destacando-se:

- a) Função gama incompleta de Wood (Wood, 1967; Wasike et al., 2011) – Largamente usada depois de sua introdução. Há somente três parâmetros que correspondem a interpretações biológicas. As curvas, no entanto, não eram suaves, mas apresentavam problemas relacionados com a forma da curva que não seguiam a função de Wood.

- b) Função exponencial de Wilmink (Wilmink, 1987) – função linear dos dias em lactação com três parâmetros. Não possui uma interpretação biológica tão clara quanto a de Wood;
- c) Função logarítmica de Ali e Schaeffer (Ali & Schaeffer, 1987) – função logarítmica que usa cinco parâmetros;
- d) Curvas não-paramétricas como a *spline* cúbica natural e B-splines (White et al., 1999; Bohmanova et al., 2008; Abdollahpour et al., 2013) – curva de lactação dividida em nós, localizados onde se assume haver mudança linear. Aqui o desafio é determinar o número de nós e sua localização ao longo da curva (Jamrozik et al. 2010).
- e) Polinômios ortogonais de Legendre (Kirkpatrick et al., 1991; Brotherstone et al., 2000, Costa et al., 2008): foi proposto o uso de funções de covariância para dados longitudinais. O usuário deve decidir a ordem que melhor ajusta a curva de lactação.

Vários estudos têm sido conduzidos com o objetivo de determinar a estrutura de covariância mais adequada e o uso de bases de dados com números diferentes de registros no dia do controle. Uma preocupação dos pesquisadores está em torno do baixo número de registros de produção no dia do controle, que pode influenciar as estimativas de valores genéticos dos animais (Pool & Meuwissen, 2001; Melo et al., 2007; Costa et al., 2005; Kim et al., 2009; Zavadilová et al. 2011; Torshizi et al., 2012). Para isso é necessário que melhores modelos de regressão sejam ajustados para assim minimizar problemas nos extremos da trajetória provocados pelo menor número de registros e alta flutuação nas variâncias e valores genéticos no início e fim dos dias em lactação (Costa et al., 2005; Melo et al., 2005; Torshizi et al., 2012).

Pool et al. (2000) aplicaram modelos de regressão aleatória com polinômios de Legendre de até sexta ordem, ou seja, com seis polinômios de Legendre para ajustar lactações completas de vacas holandesas. Os autores relataram que polinômios de quarta ordem ou superior modelaram a forma das curvas de variância ao longo dos dias em lactação com suficiente acurácia. Entretanto os mesmos autores também sugerem que um posto de três para matrix de covariância genética e quatro para matriz de ambiente permanente permite uma função de covariância mais simples com um menor número de parâmetros, baseado em resultados de autovalores e autovetores.

Um estudo com o intuito de encontrar o melhor método para prever valores genéticos aos 305 dias de vacas da raça Holandesa na Mongólia concluiu que um modelo de regressão aleatória ajustado por polinômios de Legendre de quarta ordem foram mais acurados ao utilizar mais de quatro registros de controle leiteiro por lactação (Naranchuluum et al., 2011). Zavadilová et al. (2005, 2011) utilizaram modelos de regressão aleatória e ajuste de polinômios de quarta ordem sobre dados de escore de células somáticas (logaritmo da contagem de células somáticas) contendo entre 8 a 10 registros no dia do controle por lactação para estimar parâmetros genéticos em vacas da raça Holandesa na República Tcheca. No Irã, Torshizi et al. (2012), ao utilizarem 3 a 10 registros de produção de leite no dia do controle de vacas da raça Holandesa relataram que modelos de regressão aleatória com polinômios de Legendre de terceira ordem e função matemática de Ali e Schaeffer foram

considerados ótimos e parcimoniosos modelos para o ajustamento das curvas de produção devido a limitação computacional. Ao compararem modelos com polinômios de Legendre de terceira e quarta ordem com modelo de Ali e Schaeffer apresentaram resultados semelhantes.

Mohammadi e Alijani (2014) compararam modelos de regressão aleatória ajustados por diferentes ordens de polinômios de Legendre de 2 a 6 para efeitos genético aditivo e de 2 a 6 para permanente de ambiente, usando dados de vacas com 5 registros de produção de leite, gordura e proteína no dia do controle de vacas da raça Holandesa no Irã. Nesse trabalho, foi concluído que os modelos de regressão com polinômios de Legendre para efeitos genético aditivo e para ambiente permanente foram de ordem 2 e 6 para produção de leite e de 5 e 6 para produção de gordura e proteína, respectivamente. O uso de funções diferentes para efeito genético aditivo e para efeito de ambiente permanente tem sido proposto e utilizado por vários autores como uma alternativa para ajustar modelos polinomiais de Legendre (Liu et al., 2006; Çankaya et al., 2014; Takma & Akbas, 2007, 2009).

Em alguns países, modelos de regressão aleatória já são usados nas avaliações genéticas há algum tempo. Modelos de regressão aleatória com polinômios de terceira ordem são usados na Alemanha (Liu et al., 2001), de quarta ordem no Canadá (Kistemaker, 2003) e na Itália (Muir *et al.*, 2007) e de quinta ordem no Reino Unido (Mrode *et al.*, 2003) nos programas oficiais de melhoramento dos respectivos países. No Japão, são usados modelos de regressão aleatória bicaracterística e duas lactações com no mínimo 7 controles leiteiros por vaca para avaliar touros e vacas da raça Holandesa, utilizando-se inclusive uma medida de persistência na lactação (Yamazaki et al., 2013).

No Brasil, vários estudos têm sido desenvolvidos com o intuito de determinar o melhor modelo para substituir o modelo de lactação aos 305 dias para a raça Holandesa nas condições brasileiras, sendo que na maioria deles, são consideradas como restrições, vacas com pelo menos 4 (Costa et al. 2008), 6 (Biassus et al., 2011) ou 10 (Araújo et al., 2006) registros no dia do controle. Costa et al. (2009) e Sawalha et al. (2005) compararam modelos autorregressivos como alternativos aos de produção 305 dias e concluíram que os modelos autorregressivos estimaram os valores genéticos de animais com maior acurácia que o segundo. Melo et al. (2007) comparam modelos de regressão aleatória ajustados pela função polinomial de Ali Schaeffer (AS) e a exponencial de Wilmink (W) com um modelo de repetibilidade e um modelo de lactação aos 305 dias. Esses autores concluíram que o modelo de regressão com a função AS foi o melhor ajuste alternativamente ao ajuste aos 305 dias para avaliação genética dos animais da raça Holandesa no Brasil. Araújo et al., (2006), Costa et al. (2008), Dorneles et al. (2009) e Biassus et al. (2011) indicaram modelos de regressão aleatória com polinômios de pelo menos quarta ordem são adequados para avaliações genéticas para leite, gordura e proteína de vacas da raça Holandesa no Brasil.

### 1.2.2 Comparação entre modelos

A definição dos melhores modelos de regressão aleatória para uso numaraçadependerá da utilização de testes estatísticos, estimação de parâmetros genéticos e de valores genéticos (Takma & Akbas, 2009; Aliloo et al., 2014). Os testes estatísticos mais comumente utilizados para avaliar a qualidade do ajuste dos modelos de regressão aleatória são o Critério de Informação de Akaike, AIC (Akaike, 1973), Critério de Informação Bayesiano, BIC (Schwarz, 1978), o valor da variância residual (VR) e o valor do  $-2\log L$  da função da máxima verossimilhança. Menores valores obtidos por esses critérios indicam modelos com melhor qualidade de ajuste (Faro & Albuquerque, 2005; Liu et al., 2006; Çankaya et al., 2014). Os valores de AIC, BIC e RV permitem comparar modelos não aninhados e penalizam modelos com maior número de parâmetros, sendo que o BIC atribui uma penalidade mais rigorosa (Nunez-Anton & Zimmerman, 2000; Bignardi et al., 2009).

Os testes AIC e BIC são computados como:

$$AIC = -2\log L + 2p$$

$$BIC = -2\log L + p \log(\lambda)$$

Onde  $p$  é o número de parâmetros. Usando o método REML,  $\lambda = n - r(X)$ , em que  $n$  é igual ao número de registros e  $r(X)$  igual ao posto da matriz de incidência dos efeitos sistemáticos.

Adicionalmente, são utilizados outros testes como o teste da razão da verossimilhança (LRT) cuja distribuição de qui-quadrado ( $\chi^2$ ) indica se dois modelos com diferentes números de parâmetros ou diferentes ordens de polinômios de Legendre são significativamente diferentes (Çankaya et al., 2014). A comparação dos valores observados com os valores preditos é baseada na estatística  $G$ , dada pela seguinte equação:

$$G = -2 \ln \left[ \frac{\text{verossimilhança sem avariável}}{\text{verossimilhança com avariável}} \right]$$

Ou ainda:

$$G = -2 \ln(L_s) + 2 \ln(L_c),$$

Onde  $L_s$  é a verossimilhança do modelo sem a variável e  $L_c$  é a verossimilhança do modelo com a variável. A hipótese nula que se quer testar é se a inclusão da variável, com distribuição qui-quadrado difere significativamente:



$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

A hipótese a ser testada é se os dois modelos são equivalentes ( $H_0$ ). Se o valor do teste for maior do que o valor crítico do nível de significância a hipótese será rejeitada (Tholon & Queiroz, 2011). No LRT, a diferença entre os LogL é testada, usando o qhi-quadrado com os graus de liberdade determinados como o número de parâmetros diferenças entre modelos (Huelsenbeck & Bull, 1996).

Em geral, esses testes indicam os modelos mais complexos como aqueles que melhor ajustam os registros no dia do controle ao longo dos dias em lactação, porém diferentes testes também podem indicar resultados conflitantes e tornar a escolha dos modelos uma tarefa difícil (Biassus et al., 2010; Aliloo et al., 2014). Com o intuito de diminuir problemas com resultados divergentes, Liu et al. (2006) e Aliloo et al. (2014) propuseram a utilização de um índice que reúne vários testes estatísticos num único valor, conforme equação seguinte:

$$INDEX = TRV + \log(L) + AIC + BIC + ICOMP - PRRC$$

TRV é a variância residual total, que é dada por  $TRV = \sum_{t=6}^{305} \sigma_e^2(t)$ ;  $\log(L)$  é o log da função de verossimilhança; AIC é o Critério de Informação de Akaike (Akaike, 1973); BIC é o Critério de Informação Bayesiano (Schwarz, 1978); ICOMP é a medida de informação da complexidade do modelo (Bozdogan, 2000), dado por  $ICOMP = -2\log L + C_1(\Sigma)$ ; PRRC é a porcentagem relativa da redução de complexidade, dado por  $PRRC = [(C_1(\Sigma_k) - C_1(\Sigma_{KR})) / (C_1(\Sigma_k))] \times 100\%$ ;  $\Sigma_k$  e  $\Sigma_{KR}$  são matrizes de correlação e covariância dos parâmetros do modelo k, sendo que  $C_1(\Sigma) = \text{posto}(\Sigma) \log[\text{posto}(\Sigma) / \text{posto}(\Sigma)] - \log(|\Sigma|)$ .

No entanto, esse índice também pode indicar modelos mais complexos como mais adequados para estimar as curvas ao longo da lactação (Aliloo et al., 2014). Alternativamente e/ou conjuntamente, vários pesquisadores têm utilizado autovalores e autovetores das matrizes de covariância dos efeitos aditivos genéticos e de ambiente permanente para testar a qualidade dos ajustes (Takma & Akbas, 2009; Torshizi et al., 2013; Çankaya et al., 2014). Autovalores indicam quanto da variação é explicada pela autofunção correspondente (Kirkpatrick et al., 1990). Em geral, os autovalores costumam indicar modelos com menos parâmetros como os de melhor qualidade de ajuste (Takma & Akbas, 2009; Torshizi et al., 2013; Çankaya et al., 2014).

### **1.3 HIPÓTESES**

1. A diferenciação na estrutura da base de dados por meio de restrições nos registros no dia do controle por lactação contribui para alterar a confiabilidade da predição dos valores genéticos dos animais;
2. O modelo de regressão aleatória é mais acurado que o modelo de lactação aos 305 dias para predizer o valor genético dos animais;

#### 1.4 OBJETIVOS

1. Predizer valores genéticos para subpopulações (base de dados) de animais com pelo menos 4, 6, 8 e 10 registros por lactação, ajustados pelos modelos de regressão aleatória com polinômios de Legendre de 3ª a 5ª ordem;
2. Estimar a confiabilidade dos valores genéticos de animais para produção de leite, gordura e proteína pelos modelos de lactação aos 305 dias e de regressão aleatória;



## 2. CAPÍTULO II

### SELECTING THE SUITABLE RANDOM REGRESSION MODEL FOR GENETIC EVALUATIONS OF BRAZILIAN HOLSTEIN CATTLE<sup>2</sup>

---

<sup>2</sup>Elaborado de acordo com as normas da revista *Animal*



## Selecting the suitable random regression model for genetic evaluations of Brazilian Holstein cattle

A.H. Padilha<sup>1</sup>, J.A. Cobuci<sup>1</sup>, C.N. Costa<sup>2</sup> and J. Braccini Neto<sup>1</sup>

<sup>1</sup>*Animal Science Department, Universidade Federal do Rio Grande do Sul-UFRGS, Avenida Bento Gonçalves, 7712, RS, Brazil*

<sup>2</sup>*EMBRAPA, Juiz de Fora, Minas Gerais, Brazil*

Corresponding author: Alessandro Haiduck Padilha. Email: jaime.cobuci@ufrgs.br

Short title: Selecting random regression model for Holstein cattle

### Abstract

In Brazil, genetic evaluations of Holstein cattle may be improved by replacing the current 305-day lactation model by the best random regression model. Therefore, test day milk yield records from first lactation collected by technicians of the Milk Control and Genealogy Service of the Brazilian Association of Holstein Breeders (ABCBRH) between 1991 and 2011 were used. Four subsets with different structures by restricting, in the original data, cows with at least 4, 6, 8 or 10 test days (TD) records and a maximum of 11 records in lactation, were edited. Random regression models with third (M3), fourth (M4) and fifth-order (M5) Legendre polynomial were used to fit test days of each subset. Residual variances were considered homogeneous. Single trait random regression animal models were performed by REMLF90 software. AIC, BIC and -2LogL and residual values decreased with the increasing in the polynomial orders for structure with 8 and 10 TD. For structure with

4 and 6 TD, M4 and M5 showed the lowest values. The first three or four additive genetic eigenvalues accounted for more than 99% of total variation in models M3, M4 and M5 in structures with 4, 6, 8 and 10 TD. Additive genetic correlations between test-days ranged from -0.45 to 1.00, from 0.00 to 0.99 and from -0.06 to 0.99 for models with third, fourth and fifth orders of polynomials. Heritability on days in milk ranged from 0.24 to 0.48 for M3 and from 0.17 to 0.31 for M4 and M5. Spearman correlations of EBVs of bulls and cows between M3, M4 and M5 were higher than 0.99 in all structures. Average reliability of EBVs of bulls did not increase with the increasing in the order of polynomials and were around 0.82, 0.82, 0.80 and 0.63 in structures with 4, 6, 8 and 10 TD, respectively. Differences in the number of TD in lactation influenced the reliability of EBVs for 305-day yield. A fourth-order Legendre polynomial to fit datasets with at least 4 TD is recommended for genetic evaluations of Holstein cattle in Brazil.

**Keywords:** Legendre polynomial, test day milk yield, heritability, genetic correlations

### **Implications**

The adoption of a random regression model by replacing the current 305-day lactation model also depends on the definition of the minimum number of test day records to be considered in analysis. The use of a random regression model with a fourth order Legendre polynomial is recommended and will be advantageous for Brazilian Holstein breeding program mainly because of the increasing in the number of cows and bulls considered in the genetic evaluations when cows with at least four test days are considered in the data sets.



## Introduction

The estimated breeding values (EBVs) have been used as a traditional tool for enabling the selection decision of best bulls and cows to produce the next generation in dairy cattle. Thus the increasing in the efficiency of dairy breeding schemes depends on the suitable models used for genetic evaluations in countries where genomic evaluation is still not a reality.

In many countries, the estimation of EBVs has been implemented by means of a 305-day lactation model which uses the information of the cumulative milk, fat and protein yield at 305 days (Bilal and Khan, 2009; Cobuci et al., 2011). Alternatively, the directly use of test day records for the analysis of longitudinal data, mainly by random regression models, have been officially adopted by Canada, United Kingdom and Italy, for example (Muir et al., 2007) as well as it has been recommended by many studies in literature as the best approach for genetic evaluations (Strabel et al., 2004; Dzomba et al., 2010; Khorshidie et al., 2012; Çankaya et al., 2014; Torshizi et al., 2013). In Brazil, many researchers have developed studies based on populations of Holsteins (Araújo et al., 2006; Dorneles et al., 2009; Biassus et al., 2011) or based on all Brazilian Holstein population (Costa et al., 2008) in order to found the most suitable Legendre polynomial order to be used in random regression model.

Random regression models present many advantages when compared to the lactation models as flexibility to account for the environmental and genetic components of the shape of the lactation curve (Schaeffer and Dekkers, 1994; Costa et al., 2008), possibility of the assessment of lactation persistency measures and improvement of evaluation accuracy (Pool et al., 2000; Jakobsen et al., 2002; Torshizi et al., 2013; Cobuci et al., 2011) and the evaluation of different number of test day records without the use of factors to extend lactation in progress (Bilal and

Khan, 2009). Milk yield performance is recorded at several times along lactation (longitudinal data) and each cow is expected to present about 10 monthly records during lactation. However, in literature, different restrictions have applied in datasets in order to obtain the highest accuracy of EBVs in genetic evaluations, that is, only cows with 8 (Zavadilová et al., 2011), 6 (Biassus et al., 2011) or 3 (Torshizi et al., 2013) test days are considered in the datasets, for example. In one hand, applying severe restrictions in datasets may bring a problem for genetic evaluations of Brazilian breeding schemes once Brazil still may not account for a great amount of animals recorded. On other hand, the accuracy of EBVs is crucial for enabling selection decisions, decreasing in the generation intervals and increasing the genetic gain.

Thus the purpose of the present study was to use random regression models fitted by third, fourth and fifth order of Legendre polynomials indifferent structures or subpopulations of bulls and cows by editing different subsets with Brazilian Holstein cows with 4, 6, 8 and 10 test day (TD) records in lactation in order to compare the effect of restrictions and models on the EBVs, reliability and genetic parameters.

### **Material and methods**

Data consisted of test day milk yield records collected by the technicians of the Milk Control and Genealogy Service of the Brazilian Association of Holstein Breeders (Associação Brasileira dos Criadores de Bovinos da raça Holandesa - ABCBRH) and its state affiliates between 1991 and 2011.

The original data was edited for cows aged between 18 to 48 months. The test day milk yield records were deleted if they were out of 2.5 standard-deviations. Contemporary groups of herd-year-month of test that did not contain at least four

animals and bulls with less than one daughter in two different herds were eliminated. Monthly test day milk yield records from the first lactation were used from the original data in order to prepare four subsets of the dataset (subpopulations) with different structures by restricting, in the original data, cows with at least 4, 6, 8 or 10 test day (TD) records and a maximum of 11 records in lactation, which will be designated as structure with 4, 6, 8 and 10 TD (Table 1).

After the restrictions, four classes of age at calving (18 to 25, 26 to 27, 28 to 29 and 30 to 48 months) and four calving seasons (January through March, April through June, July through September and October through December) were combined to produce 16 age-season classes.

The genetic evaluations of animals of each structure (or subpopulations) were performed by random regression models fitted by Legendre polynomials of third, fourth or fifth orders, designated as M3, M4 and M5, respectively.

The Legendre polynomials are defined for range of -1 to +1, thus the days in milk (DIM) values were transformed as below (Kirkpatrick et al., 1994),

$$DIM_t^* = -1 + \frac{2(DIM_t - DIM_{min})}{DIM_{max} - DIM_{min}}$$

Where,  $DIM_{min}$  and  $DIM_{max}$  are minimum and maximum values for the days in milk variable data.

For the t-th standardized days in milk ( $DIM_t^*$ ), the k-th polynomials is given as follows,

$$\phi(DIM_t^*)^k = \frac{1}{2^k} \sqrt{\frac{2k+1}{2}} \sum_{m=0}^{k/2} (-1)^m \binom{k}{m} \binom{2k+1}{r} (DIM_t^*)^{r-2m}$$

Where  $k/2 = (k-1)/2$  if  $k$  is odd and  $m$  is an index number needed to determine the  $k$ -th polynomial.

The random regression model used to estimate breeding values and genetic parameter were as follows:

$$y_{ijkl} = HYM_i + \sum_{k=0}^{nf} \phi_{jk} \beta_k + \sum_{k=0}^{nr} \phi_{jk} u_{jk} + \sum_{k=0}^{nr} \phi_{jk} pe_{jk} + e_{ijkl}$$

Where  $y_{ijkl}$  is the  $l$ -th test day record of cow  $j$  made on day  $t$  within HYM (herd-year-month of test) subclass  $i$ ;  $\beta_k$  are fixed regression coefficients of test-day yield that describes the regression fixed within calving age-season classes;  $HYM_i$  = fixed effect herd-year-month of testing;  $u_{jk}$  and  $pe_{jk}$  are the  $k^{\text{th}}$  random regression coefficients that describe, respectively, the additive genetic effects and the permanent environmental effects on cow  $j$ ;  $\phi_{jk}$  is the  $k^{\text{th}}$  Legendre polynomials (3<sup>rd</sup> to 5<sup>th</sup> order) for the test day record of cow  $j$  made on day  $t$ ;  $nf$  is the order of polynomials fitted as fixed regressions;  $nr$  is the order of polynomials for animal and  $pe$  effects; and  $e_{ijkl}$  is the random residual.

It was assumed that:

$$\text{var} \begin{bmatrix} \mathbf{u} \\ \mathbf{pe} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \otimes \mathbf{G} & 0 & 0 \\ 0 & \mathbf{I} \otimes \mathbf{P} & 0 \\ 0 & 0 & \mathbf{R} \end{bmatrix}$$

$\mathbf{G}$  and  $\mathbf{P}$  are covariance matrices of the random regression coefficients,  $\mathbf{R} = \mathbf{I}\sigma_e^2$  is a diagonal matrix (residual) and  $\otimes$  is a Kronecker product between matrices.

The estimates of covariance matrices of random regression coefficients and standard error prediction (SEP) were calculated by the Method of Restricted Maximum

Likelihood (REML) using the REMLF90 software (Misztal, 2014). Estimated breeding values (EBVs) of animal  $i$  on day  $t$  from random regression models was calculated by:

$$EBV_{it} = \mathbf{z}'_t \hat{\boldsymbol{\alpha}}_i = \sum_{j=0}^{k_a-1} \alpha_{ij} \phi_j(DIM_t)$$

Where  $\hat{\boldsymbol{\alpha}}_i$  is a ( $k_a \times 1$ ) vector of the estimates of additive genetic random regression coefficients specific to the animal  $i$ , and  $\mathbf{z}_t$  is a ( $k_a \times 1$ ) vector of Legendre polynomial coefficients evaluated at day  $t$ , that are presented, for example, for a fifth order polynomial:

$$\hat{\boldsymbol{\alpha}}_i = \begin{bmatrix} \hat{\alpha}_{i0} \\ \hat{\alpha}_{i1} \\ \hat{\alpha}_{i2} \\ \hat{\alpha}_{i3} \\ \hat{\alpha}_{i4} \end{bmatrix}, \quad \mathbf{z}_t = \begin{bmatrix} \phi_{0t} \\ \phi_{1t} \\ \phi_{2t} \\ \phi_{3t} \\ \phi_{4t} \end{bmatrix}$$

The sum of EBVs at 305 days for animal  $i$ , considering a fifth order polynomial, was obtained by summing the EBVs from 6 to 305 days in milk:

$$EBV_i = \sum_{t=6}^{305} (\hat{\alpha}_{i0} \phi_{0t} + \hat{\alpha}_{i1} \phi_{1t} + \hat{\alpha}_{i2} \phi_{2t} + \hat{\alpha}_{i3} \phi_{3t} + \hat{\alpha}_{i4} \phi_{4t})$$

Reliability of EBVs were derived from PEV, as  $r^2 = 1 - (SEP^2 / \sigma_a^2)$ , where  $\sigma_a^2$  was the additive genetic variance for the trait, SEP was the square root of prediction error variance (PEV) and  $r^2$  is the correlation between the true breeding value and estimated breeding values (Mistal, 2014).

Models were compared according to the goodness of fit, using the Akaike's information criterion,  $AIC = -2 \log L + 2p$ , and Schwarz's Bayesian criterion,  $BIC = -2 \log L + p \log(\lambda)$ , where  $p$  is the number of parameters in the model. Using REML,  $\lambda = n - r(X)$ ,  $n$  being equal to the number of test day records and  $r(X)$  equal to the rank of the systematic effects incidence matrix. The most suitable model is chosen based on the lowest values of AIC and BIC. Eigenvalues of additive genetic and permanent environmental covariance matrices were used to quantify contribution of higher order of Legendre polynomials. Additionally, Spearman correlations of EBVs for 305-day milk yield and reliabilities of EBVs were used to compare models.

## Results

The values of AIC, BIC and  $2\text{Log}L$  for structures with at least 4 TD were lower for M3 but RV was lower for M4 (Table 2). For structure with at least 6 TD, M5 showed the lowest values of tests, which were followed by M3 and M4. AIC, BIC,  $-2\text{Log}L$  and RV decreased with the increasing of parameters from M3 to M5 in structure with at least 8 and 10 TD. RV decreased with the increasing in the number of parameters in all structures.

The first four additive genetic eigenvalues in the structure with at least 10 TD accounted for over 99% of total variance in M5 and in M4 the first three eigenvalues accounted for 99% (Table 3). In structure with at least 8 TD, the four or five eigenvalues explained it more than 99% for M4 and M5, respectively. In structures with at least 6 and 4 TD, the first three eigenvalues accounted for over 99.3% in models M3 and M4 and in model M5 the first four accounted for 99.7%. For the eigenvalues of permanent environmental covariance matrix (Table 4), regarding all structures, in model M3 the first two eigenvalues explained 100% of total variance. In

models M4 and M5, the four or five eigenvalues, respectively, accounted for 100% of total variance.

In order to synthesize the results, the additive genetic and permanent environmental correlations between 11 selected test days from 6 to 305 days in milk were presented by means of the average and standard deviations with minimum and maximum values (Table 5). Model M3 showed similar average correlations values in structure with at least 4, 6, 8 and 10 TD with negative genetic correlations between first and last test day (-0.45 and -0.73). Average genetic correlations of models M4 and M5 were very similar in different structures (between  $0.68 \pm 0.29$  and  $0.70 \pm 0.29$ ) with negative correlations (-0.01 to -0.06) between first and last test days for M4 and close to zero but positive for M5. Permanent environmental correlations between test days were very similar, considering different structures and ranged between  $0.67 \pm 0.20$  and  $0.72 \pm 0.21$  for models M4 and M5.

Heritability on days in milk ranged from 0.24 to 0.48 for M3 model in structures with 4, 6, 8 and 10 TD, showing a U shape form for the heritability curves along days in milk (Figure 1). Models M4 and M5 presented values ranging from 0.17 to 0.31 for test days along days in milk.

Spearman rank correlations of EBVs of bulls and cows for 305-day milk yield were above 0.98 for bulls and cows, in regard to models with different order of Legendre polynomials in different structures (Table 6).

The average, standard deviations with minimum and maximum values of reliabilities of EBVs at 305 days were showed for the same group of 124 bulls in different structures or subsets of data, considering only bulls with more than 10 progenies (Table 7). In regard to the order of Legendre polynomials, average reliability of bulls did not present differences and was around  $0.82 \pm 0.12$  for structure 4 and 6, around

0.80  $\pm$  0.13 for structure 8 and around 0.64  $\pm$  0.17 for structure 10. In average, average reliability values of EBVs of bulls in structure 10 was lower than those values estimated in other structures.

## **Discussion**

The results showed that models M3, M4 and M5 applied to structure with 10 test day records had the lowest values of AIC, BIC and -2LogL, followed by the models applied to structures with 8, 6 and 4 TD (Table 2). In regard to the order of Legendre polynomials, AIC, BIC, -2LogL and RV indicated higher orders as the best fit when cows with more than 8 TD were considered in datasets. For structure with 6 TD, AIC, BIC and -2LogL showed conflicting results, indicating M5 as the best model and M3 as the second best fit but RV indicated models with more parameters. In structure with 4 TD, AIC, BIC and -2logL indicated M3 as best model but RV was lower for M4. In structures with 4, 6, 8 and 10, RV decreased in average between 11% and 15% with the increasing in the order of Legendre polynomials. In general, the increasing in the number of parameters decreased the RV around 24% to 7%, as observed in studies of Biassus et al. (2011) with Brazilian Holsteins of Minas Gerais, studies of Çankaya et al. (2014) with Jersey cows in Turkey and Takma and Akbas (2009) with Holsteins in Turkey, for example. According to studies in literature, higher order polynomials are required to fit lactation curves but it is also evident that the choice of the best model is a hard task due to conflicting results (Bohmanova et al., 2008; Biassus et al., 2011; Cobuci et al., 2011; Aliloo et al., 2014). It was expected that the increasing in the order of polynomials was followed by a decreasing in the -2logL, AIC, BIC and RV values (Takma and Akbas, 2007; Biassus et al., 2011; Aliloo et al., 2014), however it was only observed in models with at least 8 and 10 test day records (structures with 8 and 10). Biassus et al. (2011), using similar restrictions for



a population of Brazilian Holsteins, that is, similar to structure with at least 6 TD in this study, found a decreasing in the values of AIC, BIC and RV with the increasing in the order of polynomials from third to sixth.

The optimal number of regression coefficients (the order of Legendre polynomial) can also be defined as the total amount of variance explained in the correlation matrix (Kirkpatrick et al., 1990; Torshizi et al., 2013). Models with higher order polynomials have been indicated by AIC, BIC and  $-2\text{LogL}$  as best models but random regression models with third and fourth order Legendre polynomials were sufficient for modeling of test day milk yield records because the first four eigenvalues explained it more than 99.0%, considering genetic covariance matrix (Table 3 and 4). The fifth eigenvalue explained it less than 1% of additive genetic and less than 2% of permanent environmental covariance matrix. Some comparisons in literature pointed out models fitted by third order as best fit (Takma and Akbas, 2009; Torshizi et al., 2013), which suggest that lower orders to fit test days may be considered for genetic evaluations. Although the same order of polynomials was preferred for genetic and permanent environmental effects in this study, the results suggested using lower orders of polynomials for additive genetic and higher for permanent environmental variances may be used for future researches. Pool et al. (2000), using data from complete lactations of Dutch cows, determined that a fourth order was recommended for genetic evaluations with sufficient accuracy for the genetic and permanent environment effects. The authors, using eigenvalues and eigenvectors, also found that a rank of four for genetic covariance matrix and of five for permanent environment was an alternative function with a reduction in the number of parameters.

Estimates of genetic correlation between test day yields did not present differences in regard to the estimates between structures with 4, 6, 8 or 10 TD (Table 5). In model M4, correlations were higher between adjacent test days, close to 1, and decreased as the distance between test-days increased. On other hand, M3 presented high negative correlations between distant test-days because of the cubic adjustment of the respective model. M5 also presented negative correlations, close to zero, between distant test days. M4 presented positive correlations between test-days, which indicate a better adjustment of the lactation curves. Biassus et al. (2011), using data from Brazilian Holsteins from Minas Gerais State, estimated genetic correlations by models fitted by third, fourth, fifth and sixth orders ranging, from 0.12 to 1.00, from 0.02 to 1.00, from 0.11 to 1.00 and from 0.18 to 1.00 for milk yield, which except the third and sixth orders, was in agreement with this study. Cobuci et al. (2011), comparing random regression models with fourth and fifth orders of Legendre polynomial in Brazilian Holsteins, estimated higher correlations between test days than that estimated in this study.

Estimates of heritability along days in milk were equal for model M3 in structures with 4, 6, 8 and 10 TD (Figure 1). Except for structure with 10 TD, the differences in the structures with 4, 6 and 8 TD did not influence the estimates of heritability when M4 and M5 models were applied, however, the structure with 10 TD presented lower heritabilities at the beginning of lactation and higher than the estimates at the final part of lactation curves, compared to the estimates of M4 and M5. The estimated heritability on days in milk for M3 presented a U shape with the highest values at the extreme parts of lactation due to the quadratic adjustment of lactation curves. Heritability estimates from M4 and M5, in general, were higher in mid-lactation and lower mainly at the beginning of lactation. In literature, trajectories of heritability along

days in milk presented low values at extreme parts of lactation and high at mid lactation, whose values ranged from 0.15 to 0.46 for models fitted by fourth (Kim et al., 2009) and fifth (Druet et al., 2003; Costa et al., 2008) orders to fit datasets with three or four test days records. Biassus et al. (2011) and Dorneles et al. (2009), using datasets with 6 test days, found heritabilities, ranging from 0.14 to 0.31, whose trajectory increased along days in milk but with higher or lower values at the beginning and end of lactation.

The estimates of breeding values for 305-day milk yield were not influenced by the increasing in the order of polynomials or by the differentiation of structures promoted by the restrictions in datasets (Table 6). Although the trajectory of genetic parameters as heritability or the correlations between test days were influenced by the different structures, the cumulative breeding values at 305 days (305-day milk yield) and the reliabilities for a group of bulls were not influenced by the structures with different number of test days in lactation (Table 7). The increasing in the order of polynomials did not influence the reliability of EBVs for 305-day milk yield (Table 7). However, reliability of EBVs for 305-day milk yield was very similar for structures with at least 4, 6 and 8 TD in lactation but structure with at least 10 TD presented the lowest average reliability for the same group of bulls evaluated (Table 7). That result is particularly important for Brazilian breeding schemes. When the restrictions are applied in order to consider only cows with at least 10 TD (a complete lactation recorded), there is a significant decreasing in the number of bulls and cows considered in the datasets as may be observed in the Table 1, which may lead to the decreasing in the reliability of EBVs for 305-day milk yield (Table 7).

## **Conclusion**

The reliability of breeding values for 305-day milk yield of bulls was lower when only cows with complete lactations were considered in datasets, which may be attributed to the low number of cows with 10 test days in lactation in Brazilian datasets of Holstein cattle. The use of at least 4, 6 or 8 test day records did not influence the reliability of breeding values for 305-day milk yield of a group of bulls. Thus the use of records of cows with at least 4 test day records in lactation is recommended for genetic evaluations in Brazil. The increasing in the order of Legendre polynomials from third to fifth order did not increase the reliability of EBVs for 305-day milk yield of bulls. A fourth-order Legendre polynomial for additive genetic and permanent environmental is recommended for genetic evaluations for Holstein cattle in Brazil.

## **Acknowledgements**

We would like to thank the Associação Brasileira de Criadores de Bovinos da Raça Holandesa (ABCBRH) for providing the valuable data. This research was developed with the support of CENAPAD-SP (Centro Nacional de Processamento de Alto Desempenhoem São Paulo), UNICAMP / FINEP project - MCT.

## **References**

- Aliloo H, Miraie-Ashtiani SR, Shahrabak MM, Urioste JI and Sadeghi M 2014. Accounting for heterogeneity of phenotypic variance in Iranian Holstein test-day milk yield records. *Livestock Science* 167, 25-32.
- Araújo CV, Torres RA, Costa CN, Torres Filho RA, Araújo SI, Lopes PS, Regazzi AJ, Pereira CS, Cobuci JA and Sarmiento JLR 2006b. Genetic modelling of daily

- milk yield using orthogonal polynomials in random regression. *Revista Brasileira Zootecnia* 35, 967-974.
- Biassus IO, Cobuci JA, Costa CN, Rorato PRN, Neto JB and Cardoso LL 2011. Genetic parameters for production traits in primiparous Holstein cows estimated by random regression models. *Revista Brasileira de Zootecnia* 40, 85-94.
- Bilal G and Khan MS 2009. Use of test-day milk yield for genetic evaluation in dairy cattle: a review. *Pakistan Veterinary Journal* 29, 35-41.
- Bohmanova J, Miglior F, Jamrozik J, Misztal I and Sullivan PG 2008. Comparison of random regression models with Legendre polynomials and linear splines for production traits and somatic cell score of Canadian Holstein cows. *Journal of Dairy Science* 91, 3627–3638.
- Çankaya S, Takma Ç, Abaci SH and Ülker MA 2014. Comparison of some random regression models for first lactation test day milk yields in Jersey cows and estimating of genetic parameters. *Kafkas Universitesi Veteriner Fakultesi Dergisi* 20, 5-10.
- Cobuci JA, Costa CN, Neto JB and Freitas AF 2011. Genetic parameters for milk production by using random regression models with different alternatives of fixed regression modeling. *Revista Brasileira de Zootecnia* 40, 557-567.
- Costa CN, De Melo C, Packer IU, Freitas A, Teixeira N and Cobuci JA 2008. Genetic parameters for test day milk yield of first lactation Holstein cows estimated by random regression using Legendre polynomials. *Revista Brasileira de Zootecnia* 4, 602-608.
- Dorneles CKP, Cobuci JA, Rorato PRN et al. 2009. Estimation of genetic parameters for Holstein cows milk production by random regression. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia* 61, 407-412.

- Druet T, Jaffrezic F, Boichard D et al. 2003. Modeling lactation curves and estimation of genetic parameters for firstlactation test-day records of French Holstein cows. *Journal of Dairy Science* 86, 2480-2490.
- Dzomba EF, Nephawe KA, Maiwashe AN, Cote SWP, Chimonyo M, Banga CB Muller C and Dzama K 2010. Random regression test day model for analysis of dairy cattle production data in South Africa: creating the framework. *South African Journal of Animal Science* 40, 273-284.
- Jakobsen JH, Madsen P, Jensen J, Pedersen J, Christensen LG and Sorensen DA 2002. Genetic parameters for milk production and persistency for Danish Holsteins estimated in random regression models using REML. *Journal of Dairy Science* 85, 1607-1616.
- Khorshidie R, Shadparvar AA, Hossein-Zadeh NG and Shakalgurabi SJ 2012. Genetic trends for 305-day milk yield and persistency in Iranian Holsteins. *Livestock Science* 144, 211–217.
- Kim BW, Lee D, Jeon J and Lee J 2009. Estimation of genetic parameters for milk production traits using a random regression test-day model in Holstein cows in Korea. *Asian-Australasian Journal of Animal Science* 22, 923-930.
- Kirkpatrick M, Thompson R and Hill WG 1994. Estimation of covariance structure of traits during growth and aging, illustrated with lactation in dairy cattle. *Genetics Research* 64, 57-69.
- Kirkpatrick, M., D. Lofsvold, and M. Bulmer. 1990. Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics* 124:979–993
- Misztal I, Tsuruta S, Aguilar, I, Legarra A, Vitezica 2014. Manual for BLUPF90 family of programs. Available:

[http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90\\_all1.pdf](http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all1.pdf). Accessed: 30<sup>th</sup>

October, 2014.

Muir BL, Kistemaker G, Jamrozik J, Canavesi F. 2007. Genetic parameters for a multiple-trait multiple-lactation random regression test-day model in Italian Holsteins. *Journal of Dairy Science* 90, 1564-1574.

Pool MH, Janss LLG and Meuwissen THE 2000. Genetic parameters of Legendre polynomials for first parity lactation curves. *Journal of Dairy Science* 83, 2640-2649.

Schaeffer, L.R.; Dekkers, J.C.M. Random regressions in animal models for test-day production in dairy cattle. In: *World congress on genetics applied to livestock production*, 5., 1994, Guelph. Proceedings... Guelph: 1994. v.17, p.443-446.

Strabel T, Szyda J, Ptak E and Jamrozik J 2004. Comparison of random regression test-day models for Polish Black and White Cattle. *Journal of Dairy Science* 88, 3688–3699.

Takma C and Akbas 2009. Comparison of fitting performance of random regression models to test day milk yields in Holstein Friesians. *Kafkas Universitesi Veteriner Fakultesi Dergisi* 15, 261-266.

Takma Ç, Akbaş Y 2007. Estimates of genetic parameters for test day milk yields of a Holstein Friesian herd in Turkey with random regression models. *Archiv Tierzucht* 50, 327-336.

Torshizi ME, Aslamenejad AA, Nassiri MR, Farhangfar H, Solkner J, Kovac M, Meszaros G and Malovrh S 2013. Analysis of test day milk yield by random regression models and evaluation of persistency in Iranian Dairy cows. *Iranian Journal of Applied Animal Science* 3, 67-76.

Zavadilová L, Wolf J, Štípková M, Němcová E and Jamrozik J 2011. Genetic parameters for somatic cell score in the first three lactations of Czech Holstein and Fleckvieh breeds using a random regression model. *Czech Journal of Animal Science* 56, 251–260.



**Table 1** Description of databases

Trait	Structure <sup>1</sup>			
	4	6	8	10
Number of records	893,094	837,155	707,466	167,949
Number of cows	103,894	93,424	75,741	16,763
Number of animals	162,490	149,344	126,069	40,006
Year of birth	1987-2008	1987-2008	1987-2008	1987-2008
Year of calving	1990-2010	1990-2010	1990-2010	1990-2010
Number of herds	787	704	578	236
Mean age of cow at calving (months)	28.27 ± 5.29	28.16 ± 5.19	27.97 ± 5.00	27.26 ± 4.49
Average daily yield (kg)	26.07 ± 6.89	26.13 ± 6.82	26.22 ± 6.65	27.25 ± 6.37

<sup>1</sup>structures of sub-datasets (sub-populations) edited in order to contain cows with at least 4, 6, 8 and 10 and a maximum of 11 test day records per lactation

**Table 2** Values of tests of goodness of fit obtained by random regression models in different structures of datasets

Structure <sup>1</sup>	Model <sup>2</sup>	p	-2log L	AIC	BIC	RV
4	M3	13	4701875.68	4701913.68	4702053.19	7.127
	M4	21	4973974.76	4974016.76	4974261.50	6.163
6	M3	13	4613523.21	4613549.21	4613699.89	7.016
	M4	21	4640286.66	4640328.66	4640572.06	6.059
	M5	31	4317167.15	4317269.15	4317588.46	5.520
8	M3	13	3693750.98	3693788.98	3693925.51	6.909
	M4	21	3652869.22	3652935.22	3653151.13	5.973
	M5	31	3630396.04	3630498.04	3630812.21	5.430
10	M3	13	878249.93	878275.93	878362.28	7.463
	M4	21	865926.59	865968.59	866108.02	6.285
	M5	31	858107.47	858169.47	858375.20	5.606

p = number of parameters; -2LogL = logarithm of the likelihood function; AIC = Akaike's information criterion; BIC = Bayesian information criterion; RV = residual variance; <sup>1</sup> different structures of sub-datasets edited in order to contain cows with at least 4, 6, 8 and 10 and a maximum of 11 test day records per lactation; <sup>2</sup> random regression models fitted by Legendre polynomials of third to fifth order

**Table 3** Eigenvalues ( $\lambda_i$ ) of the additive genetic (co)variance matrix and the proportion of total variance (%) estimated from random regression models

Structure <sup>1</sup>	Model <sup>2</sup>	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$	n
4	M3	10.97 (55.9)	8.03 (40.9)	0.62 (3.2)			3
	M4	11.16 (80.6)	2.10 (15.2)	0.48 (3.5)	0.10 (0.7)		3
6	M3	10.68 (55.4)	7.95 (41.2)	0.65 (3.4)			3
	M4	10.88 (80.2)	2.06 (15.2)	0.53 (3.9)	0.09 (0.7)		3
	M5	10.98 (80.1)	2.06 (15.0)	0.53 (3.9)	0.10 (0.7)	0.04 (0.3)	4
8	M3	10.09 (55.2)	7.59 (41.5)	0.62 (3.4)			3
	M4	10.26 (79.0)	2.05 (15.8)	0.54 (4.1)	0.13 (1.0)		4
	M5	10.33 (79.4)	2.06 (15.8)	0.50 (3.8)	0.10 (0.8)	0.03 (0.2)	4
10	M3	10.23 (58.4)	6.97 (39.8)	0.32 (1.8)			3
	M4	10.64 (84.5)	1.52 (12.1)	0.32 (2.5)	0.11 (0.9)		3
	M5	10.78 (84.2)	1.53 (11.9)	0.32 (2.5)	0.13 (1.0)	0.04 (0.31)	4

<sup>1</sup> different structures of sub-datasets edited in order to contain cows with at least 4, 6, 8 and 10 and a maximum of 11 test day records per lactation; <sup>2</sup> random regression models fitted by Legendre polynomials of third to fifth order; n=number of eigenvalues that accounted for over 99% of total variance

**Table 4** Eigenvalues ( $\lambda_i$ ) of the permanent environmental (co)variance matrix and the proportion of total variance (%) estimated from random regression models

Structure <sup>1</sup>	Model <sup>2</sup>	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$	n
4	M3	21.15 (92.6)	1.70 (7.4)	0.0			2
	M4	21.15 (75.5)	4.29 (15.3)	1.75 (6.3)	0.82 (2.9)		4
6	M3	20.16 (92.6)	1.61 (7.3)	0.0			2
	M4	20.18 (75.2)	4.19 (15.6)	1.65 (6.2)	0.81 (3.0)		4
	M5	20.23 (73.7)	4.32 (15.7)	1.68 (6.1)	0.88 (3.2)	0.34 (1.2)	5
8	M3	18.57 (92.2)	1.57 (7.8)	0.0			2
	M4	18.64 (74.9)	3.91 (15.7)	1.59 (6.4)	0.75 (3.0)		4
	M5	18.71 (73.1)	4.05 (15.8)	1.65 (6.5)	0.85 (3.3)	0.33 (1.3)	5
10	M3	16.87 (91.9)	1.48 (8.1)	0.0			2
	M4	17.05 (74.1)	3.73 (16.2)	1.56 (6.8)	0.68 (3.0)		4
	M5	17.17 (73.0)	3.82 (15.5)	1.65 (7.0)	0.76 (3.2)	0.25 (1.1)	5

<sup>1</sup> different structures of sub-datasets edited in order to contain cows with at least 4, 6, 8 and 10 and a maximum of 11 test day records per lactation; <sup>2</sup> random regression models fitted by Legendre polynomials of third to fifth order; n=number of eigenvalues that accounted for over 99% of total variance

**Table 5** Average, minimum (min) and maximum (max) additive genetic ( $r_g$ ) and permanent environmental ( $r_{pe}$ ) correlations between test days on 6 to 305 days in milk

Structure	Model <sup>1</sup>	$r_g$			$r_{pe}$		
		Mean $\pm$ s.d.	min	max	Mean $\pm$ s.d.	min	max
4	M3	0.49 $\pm$ 0.49	-0.73	1.00	0.89 $\pm$ 0.12	0.62	1.00
	M4	0.75 $\pm$ 0.24	0.06	0.99	0.72 $\pm$ 0.21	0.24	0.99
6	M3	0.50 $\pm$ 0.48	-0.68	0.98	0.89 $\pm$ 0.12	0.63	1.00
	M4	0.70 $\pm$ 0.29	0.00	0.99	0.70 $\pm$ 0.21	0.24	0.99
	M5	0.68 $\pm$ 0.31	-0.06	0.99	0.68 $\pm$ 0.20	0.28	0.98
8	M3	0.49 $\pm$ 0.46	-0.55	0.99	0.89 $\pm$ 0.11	0.63	1.00
	M4	0.70 $\pm$ 0.27	0.08	0.99	0.70 $\pm$ 0.21	0.23	0.98
	M5	0.69 $\pm$ 0.30	-0.01	1.00	0.68 $\pm$ 0.20	0.27	0.97
10	M3	0.51 $\pm$ 0.43	-0.45	0.98	0.88 $\pm$ 0.13	0.60	1.00
	M4	0.72 $\pm$ 0.28	0.06	0.99	0.68 $\pm$ 0.21	0.21	0.98
	M5	0.70 $\pm$ 0.30	-0.02	0.99	0.65 $\pm$ 0.21	0.21	0.97

<sup>1</sup> different structures of sub-datasets edited in order to contain cows with at least 4, 6, 8 and 10 and a maximum of 11 test day records per lactation; <sup>2</sup> random regression models fitted by Legendre polynomials of third to fifth order

**Table 6** Spearman rank correlations of breeding values (bulls/cows) for 305-day milk yield between random regression models

Structure1	Models2	
4	M3	M4
		0.99/0.99
6	M3	M4
	M4	M5
		0.99/0.99
8	M3	M4
	M4	M5
		0.99/0.99
10	M3	M4
	M4	M5
		0.98/0.98
	M4	M5
		0.98/0.98
		1.00/0.99

<sup>1</sup> different structures of sub-datasets edited in order to contain cows with at least 4, 6, 8 and 10 and a maximum of 11 test day records per lactation; <sup>2</sup> random regression models fitted by Legendre polynomials of third to fifth order

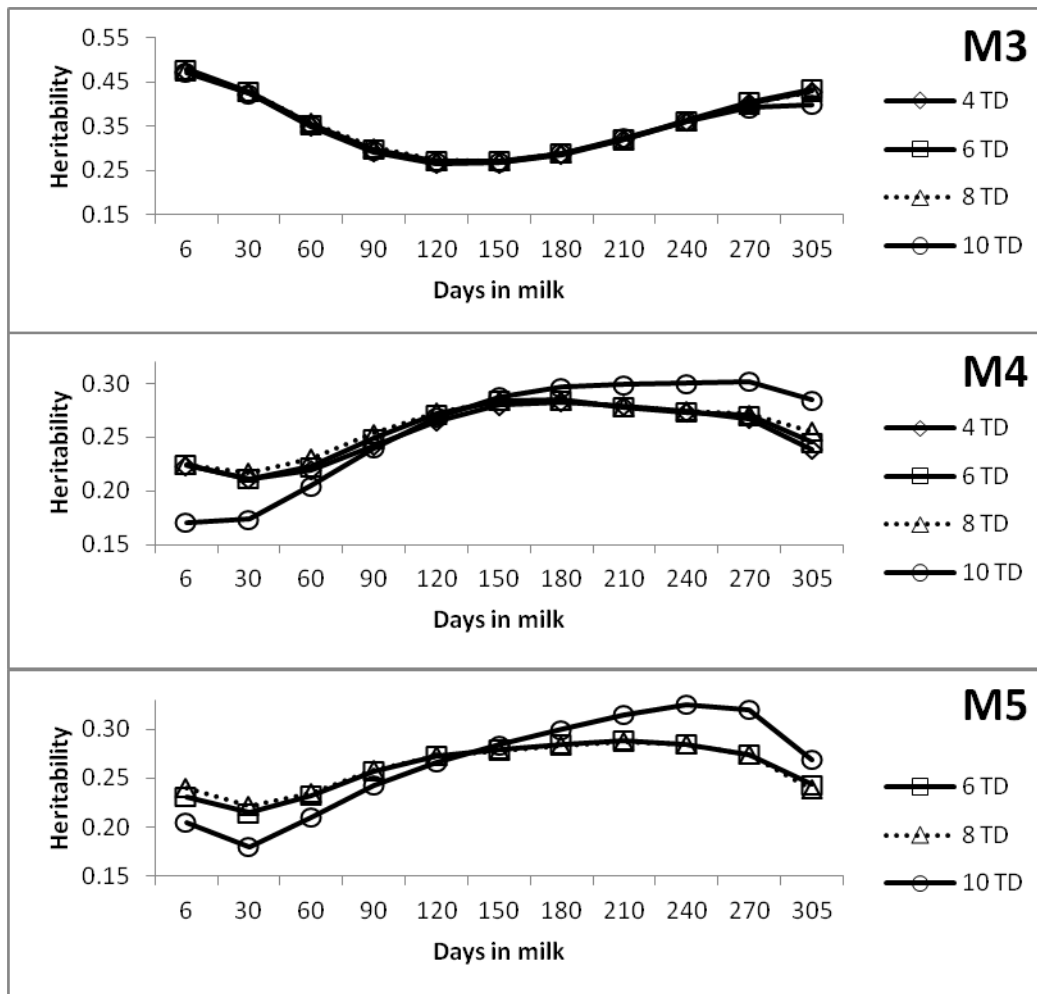
**Table 7** Average, standard deviations, minimum and maximum of reliabilities ( $r^2$ ) of estimated breeding values for 305-day yield of bulls with more than 10 progenies in different structures

Structure <sup>1</sup>	Model <sup>2</sup>	$r^2$	s.d.	Minimum	Maximum
4	M3	0.82	0.12	0.28	0.98
	M4	0.83	0.12	0.29	0.98
6	M3	0.82	0.12	0.28	0.98
	M4	0.82	0.12	0.29	0.98
	M5	0.82	0.12	0.29	0.98
8	M3	0.80	0.13	0.27	0.98
	M4	0.80	0.13	0.27	0.98
	M5	0.80	0.13	0.28	0.98
10	M3	0.63	0.17	0.12	0.94
	M4	0.64	0.17	0.12	0.94
	M5	0.64	0.17	0.12	0.94

<sup>1</sup>different structures of sub-datasets edited in order to contain cows with at least 4, 6, 8 and

10 and a maximum of 11 test day records per lactation; <sup>2</sup> random regression models fitted by

Legendre polynomials of third to fifth order



**Figure 1** Heritability of test days on days in milk estimated from random regression models with Legendre polynomials from third to fifth order (M3, M4 and M5) applied to structures with 4, 6, 8 and 10 TD



### 3. CAPÍTULO III

#### **RANDOM REGRESSION MODELS ARE SUITABLE TO SUBSTITUTE THE TRADITIONAL 305-DAY LACTATION MODEL IN GENETIC EVALUATIONS OF HOLSTEIN CATTLE IN BRAZIL<sup>3</sup>**

---

<sup>3</sup>Elaborado de acordo com as normas da revista *Czech Journal of Animal Science*





**Random regression models are suitable to substitute the traditional 305-day lactation model in genetic evaluations of Holstein cattle in Brazil**

Alessandro Haiduck Padilha<sup>1</sup>, Jaime Araujo Cobuci<sup>1</sup>, Cláudio Napolis Costa<sup>2</sup>, José Braccini Neto<sup>1</sup>

<sup>1</sup>Universidade Federal do Rio Grande do Sul, Department of Animal Science, Porto Alegre, RS, Brazil

<sup>2</sup>EMBRAPA, Juiz de Fora, MG, Brazil

**ABSTRACT:** The aim of this study was to compare two random regression models (RRM) fitted by fourth (RRM<sub>4</sub>) and fifth-order Legendre polynomials (RRM<sub>5</sub>) with a lactation model (LM) for evaluating of Holstein cattle in Brazil. Two datasets with the same restrictions and animals were prepared for this study. To apply test-day random regression models and lactation models, 262,426 test day records and 30,228 lactation records covering 305 days were prepared, respectively. The lowest values of AIC, BIC and -2LogL were for RRM<sub>4</sub>. Heritability for 305-day milk yield (305MY) was 0.23 (RRM<sub>4</sub>), 0.24 (RRM<sub>5</sub>) and 0.21 (LM). Heritability, additive genetic and permanent environmental variances of test days on days in milk was from 0.16 to 0.27, from 3.76 to 6.88 and from 11.12 to 20.21, respectively. Additive genetic and permanent environmental correlations between test days ranged from 0.20 to 0.99 and from 0.07 to 0.99, respectively. Standard deviations of average estimated breeding values (EBVs) for 305MY from RRM<sub>4</sub> and RRM<sub>5</sub> were from 11% to 30% higher for bulls and around 28% higher for cows than that in LM. Rank correlations between RRM EBVs and LM EBVs were between 0.86 to 0.96 for bulls and 0.80 to 0.87 for cows. Average percentage of gain in reliability of EBVs for 305-day yield from RRM models increased from 4% to 17% for bulls and around 28% for cows compared to LM. Random regression model fitted by

fourth order Legendre polynomials is recommended for genetic evaluations of Brazilian Holstein cattle because of the higher reliability in the estimation of breeding values.

**Keywords:** Legendre polynomials; 305-day milk yield; breeding values; reliability

## **INTRODUCTION**

In Brazil, the current estimation of breeding values (EBVs) of production traits for dairy bulls and cows in Brazilian selection programs are based on 305-day lactation model (LM). This standard approach in estimating EBVs has some disadvantages when the available test day records are projected to a 305-day lactation milk yield and the projection factors assume a standard shape of the lactation curve (Sawalha et al., 2005; Bilal and Khan, 2009). On the other hand, using test day records directly has become the most used approach in genetic models designed to breeding value estimation of production traits in dairy cattle populations, replacing the traditional 305-day lactation milk yield (305MY) evaluations (Muir et al., 2007; Dzomba et al., 2010; Bignardi et al., 2011; Cobuci et al., 2011).

The statistical modeling of test day records has many advantages and within the common used approaches, the random regression is the most appealing. Random regression models (RRM) are used to fit a linear model to obtain random regression coefficients and from them the estimation of genetic parameters and the prediction of breeding values (Jamrozik and Schaeffer, 1997). Some advantages of RRM include more accuracy for accounting environmental factors that affect cows at different stages of the lactation curve and increase of the accuracy in genetic evaluations (Costa et al., 2008, Abdullahpour et al, 2010; Naranchuluum et al. 2011). In addition, RRM permit to evaluate and rank animals (cows or bulls) for each test-day by estimating a breeding value for each test day yield. These set of breeding values define the genetic lactation curve for each animal as deviation from the

average trajectory of the population (or groups of individuals), and may allow the selection of animals whose lactation curves are more persistent (Cobuci et al, 2011).

Several studies have shown that there may be divergences between breeding values estimated by LM and RRM (Lidauer et al., 2003; Melo et al., 2007) as well as between parameters obtained by fitting RRM using different covariance functions (Kim et al., 2009; Çankaya et al., 2014). Legendre polynomials (LP) have been the preferred function to fit RRM, but there is not a consensus in literature about best order to use (Biassus et al., 2010; Çankaya et al., 2014; Aliloo et al., 2014). Canada, Italy and United Kingdom are already using a fourth or fifth order LP to fit RRM in their national genetic evaluations (Muir et al., 2007). The lack of consensus about the most suitable models enforce the importance of defining the best modeling of the covariance structure of RRM when using test day records for genetic evaluation of dairy cattle populations.

The purpose of this study was to compare genetic parameters and breeding values estimates obtained by fitting a traditional 305-day lactation model and random regression test day models using Legendre Polynomials to predict breeding values for milk yield trait of Holstein cattle in Brazil.

## **MATERIAL AND METHODS**

Data consisted of milk yield collected by the Milk Recording Services of the Brazilian Association of Holstein Breeders (ABCBRH) and its state affiliates between 1990 and 2011. The data were edited to include only records of first lactation cows between 18 and 48 months of age at calving, with full pedigree information and at least six test day records during 6 to 305 days of lactation. The records were removed from original data if 305-d lactation and test day milk yields were out of the range of 3,400 to 11,525 kg and of 9.88 to 43.45 kg, respectively.

The contemporary groups (CG) were defined by cows calving at the same herd-year-season of calving (LM) or the same herd-year-month of calving (RRM). The CG was edited to contain at least four records of cows, progeny of bulls with at least two daughters in two different herds. Four classes of age at calving (18 to 25, 26 to 27, 28 to 29 and 30 to 48 months) and four calving seasons (January through March, April through June, July through September and October through December) were combined to produce 16 age-season classes. Two datasets including the same animals, belonging to 296 herds, consisted of 30,228 lactation records totaling 262,426 test days records from daughters of 2,726 bulls were used in this study. The pedigree file included 59,486 animals.

The 305-day lactation records were analyzed by a single trait animal model (LM) including the fixed effects of herd-year-season of calving and age at calving (linear covariable) and the random additive genetic animal and residual effects in order to estimate genetic parameters and breeding values for 305-day milk yield (305MY), as follows:

$$Y_{ij} = HYS_i + b_n x_{ij} + a_{ij} + e_{ij}$$

where  $Y_{ij}$  is the cumulative milk yield at 305 days of animal  $j$  on herd-year-season of calving  $i$ ,  $HYS_i$  is the fixed effect of herd-year-season  $i$ ,  $b_n$  is the linear covariable for 305-day yield as a function of age at calving,  $x_{ij}$  is the age of cow at calving, in months;  $a_{ij}$  is the additive genetic effect of animal  $j$  on herd-year-season of calving  $i$ ,  $e_{ij}$  is the residual effect.

The test day milk yields (TDMY) were analyzed in order to obtain genetic parameters and cumulative EBVs for 305-day milk yield (305MY) using the following model:

$$y_{ijkl} = HYM_i + \sum_{k=0}^{nf} \phi_{jk} \beta_k + \sum_{k=0}^{nr} \phi_{jk} u_{jk} + \sum_{k=0}^{nr} \phi_{jk} p e_{jk} + e_{ijkl}$$

Where  $y_{ijkl}$  is the  $l$ -th test day record of cow  $j$  made on day  $t$  within HYM (herd-year-month of test) subclass  $i$ ;  $\beta_k$  are fixed regression coefficients of test-day yield that describes the regression fixed within calving age-season classes;  $\text{HYM}_i$  is fixed effect herd-year-month of testing;  $u_{jk}$  and  $pe_{jk}$  are the  $k^{\text{th}}$  random regression coefficients that describe, respectively, the additive genetic effects and the permanent environmental effects on cow  $j$ ;  $\mathcal{O}_{jtk}$  is the  $k^{\text{th}}$  Legendre polynomials (4<sup>th</sup> and 5<sup>th</sup> orders) for the test day record of cow  $j$  made on day  $t$ ;  $nf$  is the order of polynomials fitted as fixed regressions;  $nr$  is the order of polynomials for animal and  $pe$  effects; and  $e_{ijkl}$  is the random residual. The RRM referring to the fourth and fifth orders Legendre polynomials were designated as  $\text{RRM}_4$  and  $\text{RRM}_5$ , respectively. Many studies in literature have pointed out these orders as recommended as well as they have already been used in Canada, Italy and United Kingdom for genetic evaluations (Muir et al., 2007).

It was assumed that:

$$\text{var} \begin{bmatrix} \mathbf{u} \\ \mathbf{pe} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \otimes \mathbf{G} & 0 & 0 \\ 0 & \mathbf{I} \otimes \mathbf{P} & 0 \\ 0 & 0 & \mathbf{R} \end{bmatrix}$$

$\mathbf{G}$  and  $\mathbf{P}$  are covariance matrices of the random regression coefficients,  $\mathbf{R} = \mathbf{I}\sigma_e^2$  is a diagonal matrix (residual) and  $\otimes$  is a Kronecker product between matrices.

The Legendre polynomials are defined for the range of -1 to +1, thus the days in milk values were transformed as below (Kirkpatrick et al., 1994),

$$\text{DIM}_t^* = -1 + \frac{2(\text{DIM}_t - \text{DIM}_{\min})}{\text{DIM}_{\max} - \text{DIM}_{\min}}$$

Where  $\text{DIM}_{\min}$  and  $\text{DIM}_{\max}$  are minimum and maximum values for the days in milk (DIM).

For the  $t$ -th standardized days in milk ( $DIM_t^*$ ), the  $k$ -th polynomials is given as follows,

$$\phi(DIM_t^*)^k = \frac{1}{2^k} \sqrt{\frac{2k+1}{2}} \sum_{m=0}^{k/2} (-1)^m \binom{k}{m} \binom{2k+1}{r} (DIM_t^*)^{r-2m}$$

Where  $k/2 = (k-1)/2$  if  $k$  is odd and  $m$  is an index number needed to determine the  $k$ -th polynomial.

Variance components, genetic parameters and EBVs for 305-day milk yield (305MY) from RRM or LM models and the covariance matrices of random regression coefficients were estimated by using REMLF90 software (Misztal, 2014) which uses the Method of Restricted Maximum Likelihood (REML).

The EBV of animal  $i$  for test day  $t$  was calculated by:

$$EBV_{it} = \mathbf{z}'_t \hat{\boldsymbol{\alpha}}_i = \sum_{j=0}^{k_a-1} \alpha_{ij} \phi_j(DIM_t)$$

Where  $\hat{\boldsymbol{\alpha}}_i$  is a  $(k_a \times 1)$  vector of the estimates of additive genetic random regression coefficients specific to the animal  $i$ , and  $\mathbf{z}_t$  is a  $(k_a \times 1)$  vector of Legendre polynomial coefficients evaluated at day  $t$ . An example for fifth order polynomial was presented as follow:

$$\hat{\boldsymbol{\alpha}}_i = \begin{bmatrix} \hat{\alpha}_{i0} \\ \hat{\alpha}_{i1} \\ \hat{\alpha}_{i2} \\ \hat{\alpha}_{i3} \\ \hat{\alpha}_{i4} \end{bmatrix}, \quad \mathbf{z}_t = \begin{bmatrix} \phi_{0t} \\ \phi_{1t} \\ \phi_{2t} \\ \phi_{3t} \\ \phi_{4t} \end{bmatrix}$$

$\mathbf{z}'_{305}$  for fifth-order Legendre polynomial used under study was as follows:

$$\mathbf{z}'_{305} = [212.132 \ -1.276756^{-15} \ 1.586427 \ 2.529227E-15 \ 2.144969]$$

The 305-day random regression breeding value of animal  $i$  was obtained by summing the EBVs from day 6 to 305 days in milk, which was illustrated for example for a fifth order Legendre polynomial as follows,

$$EBV_i = \sum_{t=6}^{305} (\hat{\alpha}_{i0}\phi_{0t} + \hat{\alpha}_{i1}\phi_{1t} + \hat{\alpha}_{i2}\phi_{2t} + \hat{\alpha}_{i3}\phi_{3t} + \hat{\alpha}_{i4}\phi_{4t}).$$

REMLF90 software (Misztal, 2014) was also used to estimate the standard error of prediction (SEP) of EBVs. The SEP was calculated as the square root of the prediction error variance (PEV). All reliability values were derived from standard errors (SE) of the EBVs, as  $r^2 = \sqrt{1 - (SE^2/\sigma_a^2)}$ , where  $\sigma_a^2$  was the additive genetic variance for the trait and  $r^2$  is the correlation between the true breeding value and estimated breeding values (Misztal, 2005).

The models were compared according to the goodness of fit, using the Akaike's information criterion,  $AIC = -2\log L + 2p$ , and Schwarz's Bayesian criterion,  $BIC = -2\log L + p \log(\lambda)$ , where  $p$  is the number of parameters in the model. Using REML,  $\lambda = n - r(X)$ ,  $n$  being equal to the number of test day records and  $r(X)$  equal to the rank of the systematic effects incidence matrix. The model is chosen based on the lowest values of AIC and BIC. Additionally, a log-likelihood ratio test (LRT) of the likelihood was applied to test the significant differences between models with different orders of Legendre polynomials.

## RESULTS

The goodness of fit of random regression models showed lower values of  $-2\log L$ , AIC and BIC tests (Table 1) for  $RRM_4$  and lower RV for  $RRM_5$ . The difference in the  $-2\log L$  between  $RRM_4$  and  $RRM_5$  tested by LRT was significant ( $P < 0.05$ ) by the chi-square statistic.

The heritability for the 305MY estimated by RRM<sub>4</sub> and RRM<sub>5</sub> were equal 0.23 and 0.24 and slightly higher than 0.21, the value obtained for 305MY by LM (Table 2). Additive genetic variances from RRM<sub>4</sub> (402,908.3 kg<sup>2</sup>) and RRM<sub>5</sub> (400,119.7 kg<sup>2</sup>) were higher than the value from LM (311,000 kg<sup>2</sup>). The estimates of heritability, additive genetic and permanent environmental variances of test day yields from 6 to 305 days in milk ranged from 0.16 to 0.27, from 3.76 to 6.88 kg<sup>2</sup> and from 11.12 to 20.21 kg<sup>2</sup>, respectively (Figure 1).

Additive genetic and permanent environmental correlations among test days along days in milk ranged from 0.20 to 0.99 and from 0.07 to 0.99 for RRM<sub>4</sub> and RRM<sub>5</sub>, respectively (Figure 2). High additive genetic correlations were observed between adjacent test day milk yields and were close to 1 mainly during mid-lactation, but decreased with the increasing of distance between test days.

The standard deviations of average 305MY EBVs of bulls from RRM<sub>4</sub> and RRM<sub>5</sub> models increased from 11% to 31%, when progeny sizes decreased from 200-399 to 10-24 (Table 3). For cows, differences between standard deviations of 305MY EBVs from RRM<sub>4</sub> and RRM<sub>5</sub> and 305MY EBVs from LM ranged from 26% to 31%, depending on number of test days.

The rank correlation between EBVs for 305MY from LM and EBVs for 305MY from RRM<sub>4</sub> and RRM<sub>5</sub> models for bulls increased from 0.86 to 0.95 with the increase in bulls' progeny size (Table 4). Rank correlations were higher than 0.80 for cows and, in general, increased from 0.83 to 0.87 when the number of test days increased from 6 to 10.

The rank correlations between EBVs for 305MY from LM and RRM<sub>4</sub> and RRM<sub>5</sub> were equal to 0.87 and 0.86 for all cows and 0.89 (for both RRM) for all bulls (Table 5). These correlation estimates decreased from 0.89 to 0.69 when proportion of selection of top bulls was of 10% and decreased to 0.85 when proportion of selection was of 1%. When the proportion of selection was ranged from 60% to 10%, the rank correlation decreased from 0.78 to 0.57.



The selection by EBVs for 305MY from RRM<sub>4</sub> and RRM<sub>5</sub> models and the respective ranks of EBVs from LM were used to illustrate the differences between 305-day lactation and random regression models (Table 6). Figure 3 depict the trajectories of EBVs of test days on days in milk during lactation estimated by RRM<sub>4</sub> model for the top five bulls selected by RRM<sub>4</sub>.

The number of bulls within classes of reliability of EBVs estimated from RRM<sub>4</sub> and RRM<sub>5</sub> models with percentage of difference (brackets) compared to the number of bulls within classes estimated from LM was presented in Table 7. The increasing in the reliability from 0.40-0.49 to 0.90-0.99, there was an increasing in the number of bulls from 8% to 136% compared to the LM model. In lower class of reliability (0.30-0.39), there was -8% of bulls compared to LM.

The average percentage of gain in reliability of EBVs for 305MY from RRM<sub>4</sub> and RRM<sub>5</sub> models increased in average from 4% to 17% with the decrease in bulls' progeny size compared to the average of reliabilities estimated by LM, whose values ranged between 0.41 and 0.89 (Table 8). The gain of reliability in parentheses ranged from 8%-33% to 1-13% with the decrease in progeny size. For cows, the average gain in reliability was between 23% and 24% for every class of TD. Moreover, the gain in reliability of cows of each class ranged from 11-49% for cows with 6 records to 0-102% for cows with 10 test days.

## **DISCUSSION**

In a comparison between RRM<sub>4</sub> and RRM<sub>5</sub> models, AIC, BIC, -2LogL and LRT indicated RRM<sub>4</sub> as the best fit of lactation curve (Table 1). In literature, models with higher orders of Legendre polynomials were indicated as the best fit according to AIC, BIC, -2LogL and residual values (Biassus et al., 2010; Aliloo et al., 2014). Although residual variance was

lower for  $RRM_5$ , the increasing in the order of polynomials did not affect breeding values and their reliabilities as well as the estimates of genetic parameters in this study.

The heritability estimates for 305MY were equal to 0.23 and 0.24 from RRM<sub>4</sub> and RRM<sub>5</sub> and slightly higher than 0.21 obtained by the LM model but the all values showed the same magnitude (Table 2). Additive genetic variances for 305MY estimated by RRM<sub>4</sub> and RRM<sub>5</sub> were about 29% higher than that in LM model. The residual value decreased about 10% when models were fitted by fourth (RRM<sub>4</sub>) and fifth (RRM<sub>5</sub>) order Legendre polynomial. Similarly, Biassus et al. (2010) compared models fitted by Legendre polynomials from third to sixth orders whose differences decreased from 14% to 5%. Çankaya et al. (2014) compared models fitted from second to fourth orders and the results decreased from 24% to 10%. Results of residual values presented by Takma and Akbas (2009) decreased from 30% to 7% when adjusted models from second to sixth orders.

In general, heritabilities for selected test days on days in milk were higher than that estimated from LM model, except in early lactation (6, 30 and 90 days in milk) (Figure 1). Heritability estimates increased from 0.15 in the beginning up to 0.26 in mid lactation (210 days) and then decreased to 0.22 in DIM 305. Kim et al. (2009) reported higher heritability along days in milk from RRM (0.15 to 0.46) than from the LM (0.15) for Holstein cattle in Korea, which is in agreement with the present study. Araújo et al. (2006) for a population of Holstein cattle in Minas Gerais, Brazil used random regression with Legendre polynomials of third, fourth and fifth orders and estimated heritabilities 0.12 to 0.29 with an increasing in the trajectory during lactation. Costa et al. (2008) using Holstein cattle data in Brazil found heritabilities ranging from 0.27 to 0.42 with higher heritability in mid-lactation and lower at beginning and end of curves. Dorneles et al. (2009) using arandom regression model of forth order estimated heritability for Holstein cattle of Rio Grande do Sul State, Brazil, which increased from 0.14 to 0.20.

The trajectory of the additive genetic variance showed a sharp decrease in the beginning of lactation and then increased for the remaining of the lactation (Figure 1). Permanent

environmental variances from both RRM showed a typical U shape with higher variances at the extremes of lactation. This trend is similar to those reported by Miglior et al. (2009) in China and by Bignardi et al. (2009), Cobuci et al. (2011) and Pereira et al. (2013) in Brazil. Additionally, the low heritability at the beginning of lactation curve (6 to 60 days) was coincident with the lower additive genetic variance and higher permanent environmental variances at the same period. The higher heritability in mid lactation (90 to 270 days) coincided with the higher additive variances and the lower permanent environmental variances but at 305 days heritability decreased, genetic variance increased and permanent environmental variance increased.

As expected, higher genetic and permanent environmental correlations between adjacent test days (close to 1) and lower ones between distant DIM were observed in this study (Figure 2) and are in agreement with the pattern reported for Holstein cattle in Brazil (Dorneles et al., 2009; Bignardi et al. 2009; Cobuci et al., 2011). The lowest correlations were observed between early lactation (DIM 6) and the other test day milk yields. Except for DIM 6 and 30, in general, permanent environmental correlations were lower than the genetic correlations between DIM (Figure 2).

The standard deviations of EBVs for 305MY from RRM<sub>4</sub> and RRM<sub>5</sub> were higher than that estimated from LM (Table 3). When bulls' progeny size increased from higher classes of progeny size (200-299) to lower classes (10-24), the increasing in the distribution of EBVs were from 11 to 30% higher than that in the LM, which indicate that bulls with less information presented larger changes in distribution of EBVs around the mean promoted by random regression models. For cows, the change in the standard deviation of EBVs was around 28% higher in RRM models compared to LM, considering 6, 7, 8, 9 or 10 TD by lactation. Melo et al. (2007), using data from Brazilian Holsteins, found that standard deviations of EBVs from RRM was 22% higher compared to the standard deviations from

305-day lactation model for cows and the differences were up to 3% for bulls with progeny size higher than 49 and up to 22% for bulls with lower progeny size. Lidauer et al. (2003) reported an increasing about 9% for young bulls with at least 20 progenies and about 3% for active bulls with 60 progenies of Finnish dairy cattle. Therefore the increasing in the standard deviations of mean values of EBVs suggested that estimates from RRM changed the distribution of values of EBVs of bulls and cows and consequently changed the ranking of top bulls and cows.

In Table 4, the rank correlations of EBVs of bulls for 305MY estimated from LM with ranks of EBVs from RRM<sub>4</sub> and RRM<sub>5</sub> models increased from 0.86 to 0.96 according to the increasing of the progeny size classes, which indicate that the increasing in the amount of information approximate the estimation of EBVs of LM and RRM models. On other hand, as the amount of information (progeny size) decreased, the differences were higher between models, which suggest, in this case, that the re-ranking of bulls was higher for bulls with less progeny size. For cows, there was a substantial difference in the correlations between RRM models and LM in general, but, according to the number of TD, these differences were similar for cows with 6, 7, 8, 9 or 10 TD. Thus these results confirm the assumption that random regression models may change the ranking of top animals. These changes in ranking became more evident when a selection of bulls and cows were applied by RRM<sub>4</sub> and RRM<sub>5</sub> EBVs of cows and EBVs of bulls with progeny size higher than 49 (Table 5). The correlations decreased from 0.87 to 0.57 when 10% of cows were selected by RRM EBVs and increased as the proportion of selection decreased to 40% and 60% of cows. The selection of top bulls by RRM<sub>4</sub> and RRM<sub>5</sub> EBVs decreased the rank correlation from 0.89 to 0.70 (10% of bulls) and from 0.89 to 0.87 (1% of bulls). In order to illustrate the changes in the rankings of animals, the top ten bulls from the 10% of best bulls were showed (Table 6). Although the rank correlations between 10% of bulls were strong, the position of some bulls in relation to

the other bulls may have large changes as observed for bulls S4, S8 and S9. The trajectory of test-day EBVs on days in milk may show important information to explain why the cumulative EBVs for 305MY from RRM models were higher compared to EBVs from LM, which may be observed in the trajectory of the best five bulls selected for EBVs at 305 days in Figure 3. For example, bull S3 was ranked as the first best bull by LM EBVs but it did not present the best initial EBVs compared to EBVs for S1 (6 to 60 days in milk) or the best final EBVs (150 to 305 days in milk), although the EBVs in mid-lactation (60 to 150 days) were equal. Therefore the bull S1 was best than bull S3 in regard to the trajectories of TD EBVs because RRM models were able to estimate the TD EBVs (and cumulative EBVs for 305MY) with more precision and, consequently, it became more evident the differences between the two bulls, which could not have been identified by LM.

The main advantage of RRM in comparison to LM was the reliability of EBVs of bulls and cows (Table 7 and 8). There were a substantial increasing in the number of bulls whose EBVs were classified in the classes of reliability mainly for classes above 0.70-0.79, which suggested that RRM may promote an important increasing in the number of bulls with higher reliability (Table 7). Moreover, when the classes of progeny size higher than 100-199 were considered, the gain in reliability was around 4% and 5%, which showed the superiority of RRM models even when the increasing in the amount of information of bulls' progeny size were larger (Table 8). However, the gain in reliability was higher for bulls in the classes of progeny size lower than 50-99, whose percentage of gain in average were between 10% and 17%, whose ranges achieved 33%. In relation to cows, there were not differences in the average percentage of gain in reliability with the increasing in the number of test days, which may suggest that cows with lactations in progress may be evaluated earlier as well as the respective sires, decreasing the generation intervals, which, in turn, increasing the genetic

gain, with a decreasing in the costs of milk recording services for the Brazilian Holstein breeding program.

## **CONCLUSION**

Random regression models (RRM) may better explain the genetic variability of breeding values of Holsteins in Brazil and estimate these breeding values with higher reliability compared to the traditional 305-day lactation animal model. The RRM using a fourth order Legendre polynomials is recommended to be used for genetic evaluations of Brazilian Holstein cattle.

## **ACKNOWLEDGEMENTS**

We would like to thank the Associação Brasileira de Criadores de Bovinos da Raça Holandesa (ABCBRH) for providing the valuable data. This research was developed with the support of CENAPAD-SP (Centro Nacional de Processamento de Alto Desempenhoem São Paulo), UNICAMP / FINEP project - MCT.

## **REFERENCES**

- Abdollahpour R., Shahrabak, M., Nejati-Javaremi, A. and Torshizi, V.R. (2010): Genetic analysis of daily milk, fat percentage and protein percentage of Iranian first lactation Holstein cattle. *World Applied Science*, 10, 1042-1046.
- Aliloo H., Miraie-Ashtiani S.R., Shahrabak M.M., Urioste J.I., Sadeghi M. (2014): Accounting for heterogeneity of phenotypic variance in Iranian Holstein test-day milk yield records. *Livestock Science*, 167, 25-32.
- Araújo C.V., Torres R.A., Costa, C.V., Filho R.A.T., Araújo S.I., Lopes P.S., Regazzi A.J., Pereira C.P., Sarmiento J.L.R. (2006): Random regressions models to describe the genetic variation of milk yield in Holstein breed. *Revista Brasileira de Zootecnia*, 35, 975-981.

- Biassus I.O., Cobuci J.C., Costa C.N., Rorato P.R.N., Neto J.B., Cardoso L.C. (2010): Persistence in milk, fat and protein production of primiparous Holstein cows by random regression models. *Revista Brasileira de Zootecnia*, 39, 2617-2624.
- Bignardi A.B., El Faro L., Albuquerque L.G., Cardoso V.L., Machado P.F. (2009): Random regression models to estimate test day milk yield genetic parameters Holstein cows in Southeastern Brazil. *Livestock Science*, 123, 1-7.
- Bignardi A.B., Faro L., Júnior T.R.A.A., Cardoso V.L., Machado P.F., Albuquerque L.G. (2011): Random regression models using different functions to model test-day milk yield of Brazilian Holstein cows. *Genetics and Molecular Research*, 10, 4, 3565-3575.
- Bilal G., Khan M.S. (2009): Use of test-day milk yield for genetic evaluation in dairy cattle: a review. *Pakistan Veterinary Journal*, 29, 1, 35-41.
- Çankaya S., Takma Ç., Abaci S.H., Ülker M.A. (2014): Comparison of some random regression models for first lactation test day milk yields in Jersey cows and estimating of genetic parameters. *Kafkas Universitesi Veteriner Fakültesi Dergisi*, 20, 5-10.
- Cobuci J.A., Costa C.N., Neto J.B., Freitas A.F. (2011): Genetic parameters for milk production by using random regression models with different alternatives of fixed regression modeling. *Revista Brasileira de Zootecnia*, 40, 557-567.
- Costa C.N., Melo C.M.R., Packer I.U., Freitas A.F., Teixeira N.M., Cobuci J.A. (2008): Genetic parameters for test day milk yield of first lactation Holstein cows estimated by random regression using Legendre polynomials. *Revista Brasileira de Zootecnia*, 37, 602-608.
- Dorneles C.K.P., Cobuci J.A., Rorato P.R.N., Weber T., Lopes J.S., Oliveira H.N. (2009): Estimation of genetic parameters for Holstein cows milk production by random regression. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, 61, 407-412.
- Dzomba E.F., Nephawe K.A., Maiwashe A.N., Cloete S.W.P., Chimonyo M., Banga C.B., Muller C.J.C., Dzama, K. (2010): Random regression test-day model for the analysis of dairy cattle production data in South Africa: creating the framework. *South African Journal of Animal Science*, 40, 273-284.



- Jamrozik J., Schaeffer L.R. (1997) Estimates of genetic parameters for a test day model with random regressions for yield traits of first lactation Holsteins. *Journal of Dairy Science*, 80, 762-770.
- Kim B.W., Lee D., Jeon J., Lee J. (2009): Estimation of genetic parameters for milk production traits using a random regression test-day model in Holstein cows in Korea. *Asian-Australasian Journal of Animal Science*, 22, 923-930.
- Kirkpatrick M., Thompson R., Hill W.G. (1994): Estimation of covariance structure of traits during growth and aging, illustrated with lactation in dairy cattle. *Genetics Research*, 64, 57-69.
- Lidauer M., Mäntysaari E.A., Strandén I. (2003): Comparison of test-day models for genetic evaluation of production traits in dairy cattle. *Livestock Production Science*, 79, 73-86.
- Melo C.M.R., Packer I.U., Machado P.F., Patrício M. (2007): Breeding values prediction for test day and lactation milk yields in Holstein cattle using different statistical models. *Revista Brasileira de Zootecnia*, 36, 1295-1303.
- Miglior F., Gong W., Wang Y., Kistemaker G.J., Sewalem A., Jamrozik J. (2009): Short communication: Genetic parameters of production traits in Chinese Holsteins using a random regression test-day model. *Journal of Dairy Science*, 92, 4697-4706.
- Misztal I., Tsuruta S., Aguilar I., Legarra A., Vitezica Z. (2014): Manual for BLUPF90 family of programs. Available: [http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90\\_all1.pdf](http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all1.pdf). Accessed: 30 October, 2014.
- Muir B.L., Kistemaker G., Jamrozik J., Canavesi F. (2007): Genetic parameters for a multiple-trait multiple-lactation random regression test-day model in Italian Holsteins. *Journal of Dairy Science*, 90, 1564-1574.
- Naranchuluum G., Ohmiya, H., Masuda Y., Hagiya K. and Suzuki M. (2011): Selecting the desirable method for predicting 305-day lactation yields in Mongolia. *Animal Science Journal*, 82, 383-389.
- Pereira R.J., Bignardi A.B., El Faro L., Verneque R.S., VercesiFilho A.E., Albuquerque L.G. (2013) Random regression models using Legendre polynomials or linear splines for test-day milk yield of dairy Gyr (*Bos indicus*) cattle. *Journal of Dairy Science*, 96, 565-574.

Sawalha R.M., Keown J.F., Kachman S.D., Van Vleck L.D. (2005): Genetic evaluation of dairy cattle with test-day models with autoregressive covariance structures and with a 305-d model. *Journal of Dairy Science*, 88, 3346–3353.

Takma C. and Akbas Y. (2009): Comparison of fitting performance of random regression models to test day milk yields in Holstein Friesians. *Kafkas Universitesi Veteriner Fakultesi Dergisi*, 15, 261-266.

Corresponding author: JAIME ARAUJO COBUCCI, Departamento de Zootecnia, Faculdade de Agronomia, Universidade Federal do Rio Grande do Sul, Avenida Bento Gonçalves, 7712, 91540-000, Porto Alegre, RS, Brasil. E-mail: [jaimcobuci@ufrgs.br](mailto:jaimcobuci@ufrgs.br)

Table 1. Number of parameters (p), estimates of the maximum of the likelihood function (-2 Log L), Akaike's information criterion (AIC), Bayesian information criterion (BIC), residual value (RV), likelihood ratio test (LRT) and chi-square statistics ( $\chi^2$ ) for random regression models using Legendre polynomials.

Model	p	-2 Log L	AIC	BIC	RV	LRT	$\chi^2$
RRM <sub>4</sub>	21	<b>1,328,860.6</b>	<b>1,328,902.6</b>	<b>1,328,974.4</b>	5.402	-	-
RRM <sub>5</sub>	31	1,475,279.4	1,475,341.4	1,475,447.4	<b>4.887</b>	146,418.8*	18.30

\* ( $P < 0.05$ ); RRM<sub>4</sub> and RRM<sub>5</sub> = random regression models fitted by fourth and fifth order Legendre polynomials

Table 2. Estimates of additive genetic ( $\sigma^2_a$ ), residual ( $\sigma^2_e$ ) and permanent environmental ( $\sigma^2_{pe}$ ) variance components and heritability coefficients ( $h^2$ ) for 305-day milk yield (305MY) estimated from LM, RRM<sub>4</sub> and RRM<sub>5</sub> models

Model	$h^2$	$\sigma^2_a$	$\sigma^2_e$	$\sigma^2_{pe}$
RRM <sub>4</sub>	0.23	402,908.3	486,000.0	843,284.2
RRM <sub>5</sub>	0.24	400,119.7	439,830.0	849,730.5
LM	0.21	311,000.0	1,181,000	-

LM = 305-day lactation model; RRM<sub>4</sub> and RRM<sub>5</sub> = random regression models fitted by fourth and fifth order Legendre polynomials

Table 3. Standard deviations (kg) of EBVs for 305MY estimated by RRM<sub>4</sub> and RRM<sub>5</sub> models, and percentage of change (between brackets) with respect to standard deviations by LM model for bulls and cows according to progeny size and number of test days

Bulls		Models		
Progeny size	Number of bulls	LM	RRM <sub>4</sub>	RRM <sub>5</sub>
200 – 399	29	424.7	462.3 (+8%)	467.1 (+10%)
100 – 199	74	476.8	530.9(+11%)	534.0 (+12%)
50 – 99	154	412.6	481.2(+16%)	484.7(+17%)
25 – 49	175	380.3	446.2(+17%)	446.6(+17%)
10 – 24	352	323.9	421.1(+30%)	424.0 (+31%)
Cows		Models		
Number of test days	Number of cows	LM	RRM <sub>4</sub>	RRM <sub>5</sub>
10	9449	308.6	392.6 (+27%)	392.7 (+27%)
9	9569	309.1	395.5 (+28%)	396.2 (+28%)
8	5433	308.8	404.3 (+31%)	405.9 (+31%)
7	3534	327.8	418.8 (+28%)	420.8 (+28%)
6	2243	333.7	419.8 (+27%)	421.8 (+26%)

LM = 305-day lactation model; RRM<sub>4</sub> and RRM<sub>5</sub> = random regression models fitted by fourth and fifth order

Legendre polynomials; 305MY = 305-day milk yield

Table 4. Spearman rank correlation between EBVs for 305MY estimated from LM and EBVs for 305MY from RRM<sub>4</sub> and RRM<sub>5</sub> models for bulls and cows according to progeny size and number of test days

Bulls		Models	
Progeny size	Number of bulls	RRM <sub>4</sub>	RRM <sub>5</sub>
200 – 399	29	0.95	0.95
100 – 199	74	0.97	0.97
50 – 99	154	0.92	0.92
25 – 49	175	0.92	0.92
10 – 24	352	0.86	0.86
Cows		Models	
Number of test days	Number of cows	RRM <sub>4</sub>	RRM <sub>5</sub>
10	9449	0.87	0.86
9	9569	0.86	0.86
8	5433	0.86	0.85
7	3534	0.86	0.86
6	2243	0.83	0.83

LM = 305-day lactation model; RRM<sub>4</sub> and RRM<sub>5</sub> = random regression models fitted by fourth and fifth order

Legendre polynomials; 305MY = 305-day milk yield

Table 5. Spearman rank correlation ( $p < 0.0001$ ) between EBVs for 305MY estimated from LM and EBVs for 305MY from RRM<sub>4</sub> and RRM<sub>5</sub> models for bulls with progeny size higher than 49 and cows selected for 305MY

Animal selected	Number of		
		RRM <sub>4</sub>	RRM <sub>5</sub>
All Bulls	2,726	0.89	0.89
10% bulls	273	0.70	0.69
1% bulls	27	0.87	0.85
Animal selected	Number of		
		RRM <sub>4</sub>	RRM <sub>5</sub>
All cows	56,760	0.87	0.86
60% cows	34056	0.78	0.78
40% cows	22704	0.71	0.70
10% cows	5676	0.57	0.54

LM = 305-day lactation model; RRM<sub>4</sub> and RRM<sub>5</sub> = random regression models fitted by fourth and fifth order

Legendre polynomials; 305MY = 305-day milk yield

Table 6. Ranks of EBVs of the ten top bulls with progeny size higher than 49 selected by EBVs for 305MY estimated from RRM<sub>4</sub> and RRM<sub>5</sub> and respective ranks from LM

Bull	Progenysize	RRM <sub>4</sub>	RRM <sub>5</sub>	LM
S1	143	1	1	2
S2	64	2	2	3
S3	51	3	3	1
S4	80	4	4	16
S5	90	5	5	4
S6	235	6	6	5
S7	145	7	7	6
S8	162	8	8	20
S9	154	9	9	19
10	79	10	13	9

LM = 305-day lactation model; RRM<sub>4</sub> and RRM<sub>5</sub> = random regression models fitted by fourth and fifth order

Legendre polynomials; 305MY = 305-day milk yield

Table 7. Classes of reliability ( $r^2$ ) of EBVs and the number of bulls for LM and for RRM<sub>4</sub> and RRM<sub>5</sub> models with percentage of increasing in the number of bulls in the respective class (brackets) compared to LM

$r^2$	LM	RRM <sub>4</sub>	RRM <sub>5</sub>
0.90 – 0.99	22	52 (+136%)	52 (+136.0%)
0.80 – 0.89	78	123 (+58.0%)	123 (+58.0%)
0.70 – 0.79	104	122 (+17.0%)	123 (+18.0%)
0.60 – 0.69	111	130 (+17.0%)	128 (+15.0%)
0.50 – 0.59	136	160 (+18.0%)	161 (+18.0%)
0.40 – 0.49	186	202 (+9.0%)	200 (+8.0%)
0.30 – 0.39	363	335 (-8.0%)	333 (-8.0%)

LM = 305-day lactation model; RRM<sub>4</sub> and RRM<sub>5</sub> = random regression models fitted by fourth and fifth order Legendre polynomials

Table 8. Reliabilities and their standard deviations of EBVs for 305MY estimated from LM and the percentage of gain with range in brackets by RRM<sub>4</sub> and RRM<sub>5</sub> models for bulls and cows according progeny size and number of test days

Bulls		Models		
Progeny size	Number of bulls	LM <sup>1</sup>	RRM <sub>4</sub> <sup>2</sup>	RRM <sub>5</sub> <sup>2</sup>
200 – 399	29	0.89 ± 0.06	+4% (1-13%)	+4% (1-13%)
100 – 199	74	0.84 ± 0.05	+5% (2-14%)	+5% (2-14%)
50 – 99	154	0.71 ± 0.11	+10% (4-24%)	+10% (4-24%)
25 – 49	175	0.59 ± 0.10	+13% (4-25%)	+13% (4-25%)
10 - 24	352	0.41 ± 0.12	+17 % (8-33%)	+17% (8-33%)
Cows		Models		
Number of test days	Number of cows	LM	RRM <sub>4</sub>	RRM <sub>5</sub>
10	9449	0.34 ± 0.08	+24% (0-102)	+24% (0-102)
9	9569	0.34 ± 0.08	+24% (13-50)	+24% (13-51)
8	5433	0.32 ± 0.09	+24% (8-65)	+24% (9-64)
7	3534	0.31 ± 0.08	+23% (11-48)	+23% (11-48)
6	2243	0.31 ± 0.09	+23% (11-46)	+23% (11-46)

LM = 305-day lactation model; RRM<sub>4</sub> and RRM<sub>5</sub> = random regression models fitted by fourth and fifth order

Legendre polynomials; 305MY = 305-day milk yield; <sup>1</sup> average and standard-deviation of reliability of EBVs from LM; <sup>2</sup> average percentage of gain (%) in reliability compared to lactation model (range in parenthesis)

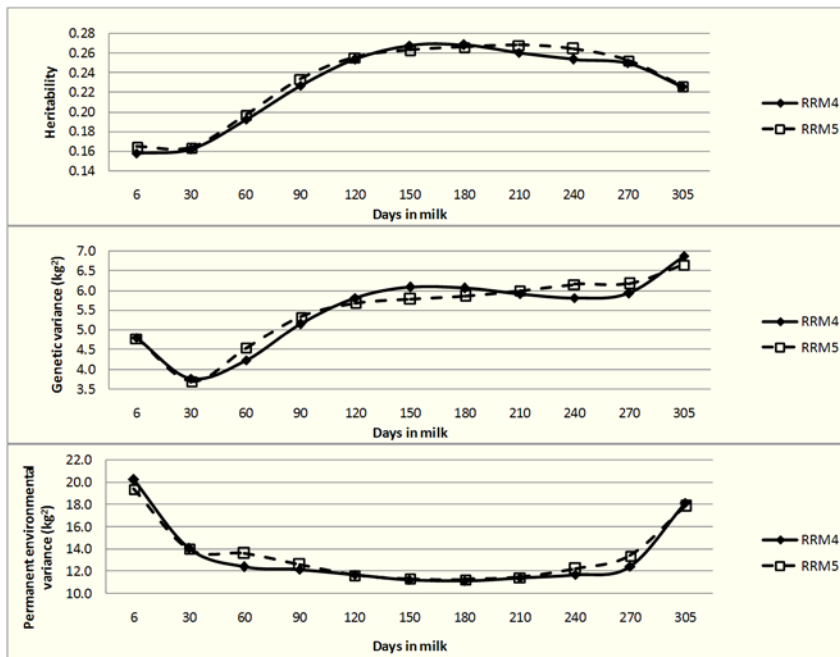


Figure 1 - Heritability, additive genetic and permanent environmental variances of test day milk yields on days in milk estimated from random regression models fitted by fourth (RRM<sub>4</sub>) and fifth (RRM<sub>5</sub>) order Legendre polynomials

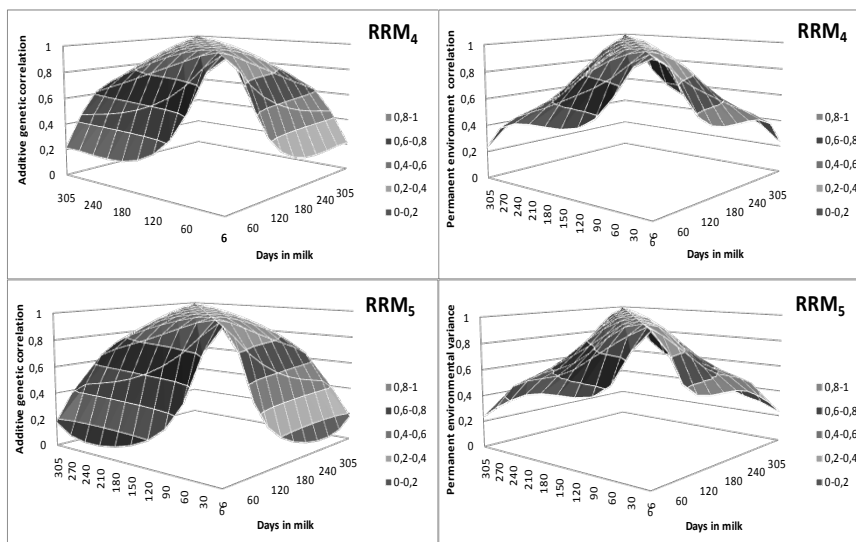


Figure 2 – Genetic correlation estimates (left) and permanent environmental correlation (right) between test-day milk yield along days in milkestimated from random regression models fitted by fourth (RRM<sub>4</sub>) and fifth (RRM<sub>5</sub>) order Legendre polynomials

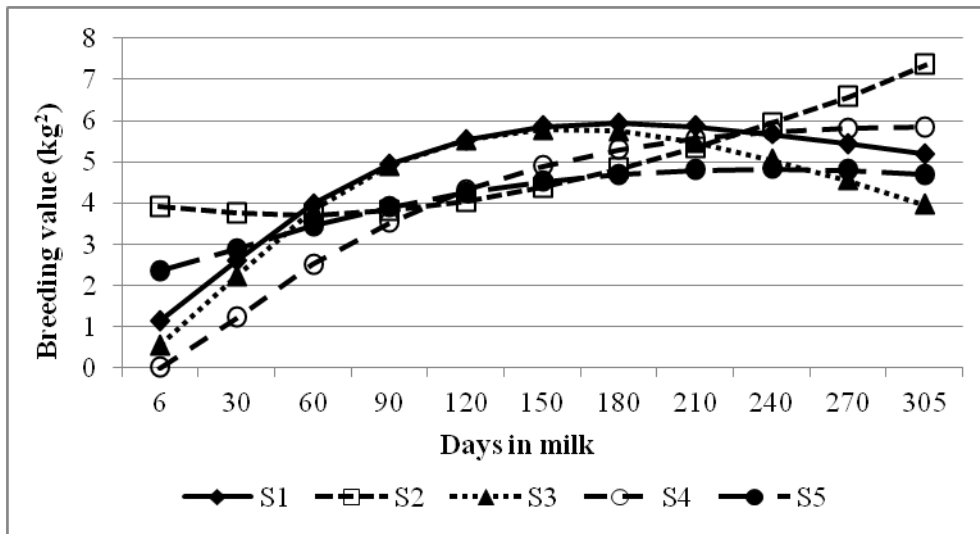


Figure 3 - Trajectory of test day breeding values ( $\text{kg}^2$ ) on days in milk of five top bulls (S1 to S5) selected on breeding values for 305day milk yield estimated from LM and  $\text{RRM}_4$  models





#### **4. CAPÍTULO IV**

### **RANDOM REGRESSION AND 305-DAY LACTATION MODELS FOR FAT AND PROTEIN YIELDS IN GENETIC EVALUATIONS OF HOLSTEIN CATTLE IN BRAZIL<sup>4</sup>**

---

<sup>4</sup>Elaborado de acordo com as normas da revista *Animal Science Journal*



**Random regression and 305-day lactation models for fat and protein yields in genetic evaluations of Holstein cattle in Brazil**

Alessandro Haiduck Padilha<sup>1</sup>, Jaime Araujo Cobuci<sup>1</sup>, Cláudio Napolis Costa <sup>2</sup>,  
José Braccini Neto<sup>1</sup>

<sup>1</sup>Universidade Federal do Rio Grande do Sul - UFRGS, Bento Gonçalves  
Avenue, 7712, Porto Alegre, RS, Brazil, CEP: 91540-000,  
[jaime.cobuci@ufrgs.br](mailto:jaime.cobuci@ufrgs.br), Phone number: +55 (51) 33087421

<sup>2</sup> EMBRAPA, Juiz de Fora, MG, Brazil, [claudio.napolis@embrapa.br](mailto:claudio.napolis@embrapa.br)

Random regression and lactation models

## **ABSTRACT**

The aim of this study was to compare random regression models fitted by fourth and fifth Legendre polynomials with lactation models using fat and protein yield records of Brazilian Holstein cattle. Two datasets were prepared. The first contained 262,426 test-day fat and protein yield records and second one contained 30,228 fat and protein yield lactation records at 305 days. Single trait random regression models with Legendre polynomials and single trait 305-day lactation models were applied. AIC, BIC, -2LogL values were lower for fourth order polynomial. Heritability for 305-day yield from lactation models was of 0.24 (fat) and 0.17 (protein) and from random regression models was of 0.20 (fat) and 0.21 (protein). Spearman correlations of EBVs for 305-day yield between lactation models and random regression models ranged from 0.86 and 0.97 for bulls and cows. In average, reliability of estimated breeding values (EBVs) for 305-day yield of bulls were from 2% to 16% (fat) and from 4% to 26% (protein) and of cows was from 24% to 38% (fat and protein) higher than in the lactation models. Random regression model fitted by Legendre polynomials of fourth order was recommended to improve genetic evaluations of Brazilian Holstein cattle.

**Key words:** Legendre polynomial, breeding value, correlation

## **INTRODUCTION**

One of the main objectives in dairy cattle breeding programs is the improvement of milk yield traits as fat and protein yield. In Brazil, the genetic evaluations of Holstein cattle for these yields have been realized by means of a 305-day

lactation model. Alternatively, many other approaches proposed using the test day records directly in test day models (TDM) instead of lactation models (Jensen, 2001).

The main advantages of the alternative approaches are that test day models permit to model the shape of the lactation curve (Schaeffer et al., 2000) and account for environmental factors that affect cows at different stages of lactation with more accuracy (Jensen, 2001). Within test day models, random regression models have been proposed for genetic evaluations on traits measured over time (Henderson, 1982; Schaeffer and Dekkers, 1994; Kirkpatrick et al., 1994), using mainly Legendre polynomials to adjust lactation curves and, then the differences between animals could be modeled as deviations of fixed lactation curves (Brothstone et al., 2000). Germany, Canada, United Kingdom and Italy have already adopted random regression models in their national genetic evaluations, using Legendre polynomials of third, fourth or fifth orders (Muir et al., 2007; Yamazaki et al., 2013). In Brazil, Costa et al. (2008), Biassus et al., (2011) and Cobuci et al. (2011), for example, studied the use of random regression models with Legendre polynomials in order to determine the best order for genetic evaluation of Holstein cattle in Brazil and then substitute the current 305-day lactation model. However, there are few studies with components of milk as fat and protein in tropical countries as Brazil, using random regression models. Costa et al. (2008), using milk yield records, suggested the use of a five Legendre polynomial order to model both genetic and permanent environmental effects. Biassus et al. (2011) used the data of milk, fat and protein yield of the population of Holstein of Minas Gerais State

and suggested that, at least, a fourth order was recommended for genetic evaluations.

The main objective was to compare genetic parameters and estimated breeding values (EBV) for fat and protein yields based on traditional 305-day lactation model and on random regression models in Brazilian Holstein cattle to be used in future genetic evaluations.

## **MATERIAL AND METHODS**

Data consisted of fat and protein milk yield collected by the technicians of the Milk Control and Genealogy Service of the Brazilian Association of Holstein Breeders (ABCBRH) and its state affiliates between 1990 and 2011.

Two datasets with the same animals were prepared. Test-day records obtained between 6 and 305 days in milk were edited for cows aged 18 to 48 months in the first parity. To apply the random regression models, 262,426 test-day fat (TDF) and protein yield (TDP) records were prepared for first data set, which comprised 6 test days in the lactation period. Test-day records were removed if fat and protein yields were out of the range of 258.4 g to 1,510 g and of 312.0 g to 1,314.8 g, respectively. Four classes of age at calving (18 to 25, 26 to 27, 28 to 29 and 30 to 48 months) and four calving seasons (January through March, April through June, July through September and October through December) were combined to produce 16 age-season classes. The second data set comprised of 30,228 fat and protein lactation records at 305 days. The records for lactation fat and protein yields at 305 days were deleted if out of the range of 102 to 392 kg and 106 to 349 kg, respectively. For the two datasets,

contemporary groups of herd-year-season of calving (lactation records) and herd-year-month of test (test day records) that did not contain at least four records of cows, progeny of bulls with at least two daughters in two different herds were eliminated. Pedigree data was checked for inconsistency and, after edition, included 59,486 animals.

Milk traits considered in the present analyses were 305-day fat yield (305F) and 305-day protein yield (305P).

Legendre polynomials are defined for the range of -1 to +1, thus the days in milk values were transformed as below,

$$d_t^* = -1 + \frac{2(d_t - d_{min})}{d_{max} - d_{min}},$$

Where  $d_{min}$  and  $d_{max}$  are minimum and maximum values for the days in milk variable data.

For the  $t$ -th standardized days in milk ( $d_t^*$ ), the  $k$ -th polynomials is given as follows,

$$\phi(d_t^*)^k = \frac{1}{2^k} \sqrt{\frac{2k+1}{2}} \sum_{m=0}^{k/2} (-1)^m \binom{k}{m} \binom{2k+1}{r} (d_t^*)^{r-2m}$$

Where  $k/2 = (k-1)/2$  if  $k$  is odd and  $m$  is an index number needed to determine the  $k$ -th polynomial.

Test-day fat and protein yield records were used in single trait test day random regression models to estimate EBVs for TDF and TDP from 6 to 305 days in



milk and the cumulative EBVs for 305F and 305P. Models were named as RRF4 and RRF5 for fat yield and RRP4 and RRP5 for protein yield fitted by fourth and fifth orders, respectively. Many studies in literature have pointed out these orders as recommended as well as they have already been used in Canada, Italy and United Kingdom for genetic evaluations (Muir et al., 2007).

The random regression model used to estimate genetic parameters and EBVs for 305F and 305P was as follows:

$$y_{ijkl} = HYM_i + \sum_{k=0}^{nf} \phi_{jk} \beta_k + \sum_{k=0}^{nr} \phi_{jk} u_{jk} + \sum_{k=0}^{nr} \phi_{jk} pe_{jk} + e_{ijkl}$$

Where  $y_{ijkl}$  is the  $l$ -th test day record of cow  $j$  made on day  $t$  within HYM (herd-year-month of test) subclass  $i$ ;  $\beta_k$  are fixed regression coefficients of test-day yield that describes the regression fixed within calving age-season classes;  $HYM_i$  = fixed effect herd-year-month of testing;  $u_{jk}$  and  $pe_{jk}$  are the  $k^{\text{th}}$  random regression coefficients that describe, respectively, the additive genetic effects and the permanent environmental effects on cow  $j$ ;  $\phi_{jk}$  is the  $k^{\text{th}}$  Legendre polynomials (4<sup>th</sup> and 5<sup>th</sup> orders) for the test day record of cow  $j$  made on day  $t$ ;  $nf$  is the order of polynomials fitted as fixed regressions;  $nr$  is the order of polynomials for animal and  $pe$  effects; and  $e_{ijkl}$  is the random residual.

It was assumed that:

$$\text{var} \begin{bmatrix} \mathbf{u} \\ \mathbf{pe} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{A} \otimes \mathbf{G} & 0 & 0 \\ 0 & \mathbf{I} \otimes \mathbf{P} & 0 \\ 0 & 0 & \mathbf{R} \end{bmatrix}$$

**G** and **P** are covariance matrices of the random regression coefficients,  $\mathbf{R} = \mathbf{I}\sigma_e^2$  is a diagonal matrix (residual) and  $\otimes$  is a Kronecker product between matrices. Records of 305-day lactation fat and protein yields were used in single trait lactation animal models, named as LMF (fat) and LMP (protein), which included effects of herd-year-season of calving as fixed effect, age at calving (linear covariable) and additive genetic animal and residual effects. The model used to estimate genetic parameters and EBVs for 305F and 305P was as follows:

$$Y_{ij} = HYS_i + b_n x_{ij} + a_{ij} + e_{ij}$$

Where  $Y_{ij}$  is the fat or protein milk yield records at 305 days of animal  $j$  on herd-year-season of calving  $i$ ,  $HYS_i$  is the fixed effect of herd-year-season of calving  $i$ ,  $b_n$  is the linear covariable for 305-day yield as a function of age at calving,  $x_{ij}$  is the age of cow at calving, in months;  $a_{ij}$  is the additive genetic effect of animal  $j$  on herd-year-season of calving  $i$ ,  $e_{ij}$  is the residual effect.

The analyses were performed by REMLF90 software (Misztal et al., 2014), by the method of restricted maximum likelihood (REML), in order to estimate the solutions and the covariance matrices of random regression coefficients.

The estimated breeding values (EBVs) of random regression models were obtained by multiplying covariance matrices and vectors containing covariates specific for each animal. The EBV of animal  $i$  for test day  $t$  was calculated by:

$$EBV_{it} = \mathbf{z}'_t \hat{\boldsymbol{\alpha}}_i = \sum_{j=0}^{k_a-1} \alpha_{ij} \phi_j(d_t)$$

Where  $\hat{\alpha}_i$  is a ( $k_a \times 1$ ) vector of the estimates of additive genetic random regression coefficients specific to the animal  $i$ , and  $\mathbf{z}_t$  is a ( $k_a \times 1$ ) vector of Legendre polynomial coefficients evaluated at day  $t$ , which may be illustrated for a fifth-order Legendre polynomial:

$$\hat{\alpha}_i = \begin{bmatrix} \hat{\alpha}_{i0} \\ \hat{\alpha}_{i1} \\ \hat{\alpha}_{i2} \\ \hat{\alpha}_{i3} \\ \hat{\alpha}_{i4} \end{bmatrix}, \quad \mathbf{z}_t = \begin{bmatrix} \phi_{0t} \\ \phi_{1t} \\ \phi_{2t} \\ \phi_{3t} \\ \phi_{4t} \end{bmatrix}$$

The sum of EBVs at 305 days for animal  $i$  was obtained by summing the EBVs from day 6 to 305:

$$EBV_i = \sum_{t=6}^{305} (\hat{\alpha}_{i0}\phi_{0t} + \hat{\alpha}_{i1}\phi_{1t} + \hat{\alpha}_{i2}\phi_{2t} + \hat{\alpha}_{i3}\phi_{3t} + \hat{\alpha}_{i4}\phi_{4t})$$

The standard error prediction (SEP) of estimated breeding values (EBVs) for 305F and 305P estimated from 305-day lactation and random regression models was supplied by REMLF90 software as the square root of the prediction error variance (PEV) (Misztal, 2014). Reliability of EBVs were derived from SEP, as  $r^2 = 1 - (\text{SEP}^2 / \sigma_a^2)$ , where  $\sigma_a^2$  was the additive genetic variance for the trait and  $r^2$  is the correlation between the true breeding value and estimated breeding values (Misztal & Wiggans, 1988).

## RESULTS

The goodness of fit for random regression models was given by AIC, BIC, -2LogL and likelihood ratio test (LRT) in Table 1. AIC, BIC and -2LogL had the lowest values (highlighted) for RRF4 and RRP4 models. Residual values showed a decreasing of 4% (fat) and 5% (protein) with an increasing in the order of Legendre polynomials. According to LRT, the changes in log likelihood were significant ( $p < 0.05$ ) in the increasing in the order of Legendre polynomials from RRF4 to RRF5 and RRP4 to RRP5 models.

The first eigenvalue ( $\lambda$ ) accounted for 89.5% in RRF4 and RRF5 and for 88.7 and 89.0% in RRP4 and RRP5 of the total additive genetic covariance matrix, and the first three eigenvalues explained it by 99.99% (Table 2). For permanent environmental effect, the first eigenvalue accounted for about 70% of the total permanent environmental variance in all models and the first four eigenvalues explained it by 98%.

For 305-day yield trait (Table 3), heritability from 305-day lactation models were of 0.24 (LMF) and 0.17 (LMP) and random regression models presented heritability of 0.21 (RRF4 and RRF5) and 0.20 (RRP4 and RRP5).

In regard to heritability of test-day yields on days in milk, (RRF4 and RRF5 models) showed an increasing from 0.13 (6 days in milk) to 0.23 (210 days in milk) and then decreased to 0.14 (305 days in milk). RRP4 and RRP5 models showed similar trajectory of heritability with an increasing from 0.10 (6 days in milk) to 0.23 (210 days in milk) and then decreased to 0.13 (Figure 1). Additive genetic variances of test days during days in milk ranged from 4,954.4  $g^2$  to 7,115.6  $g^2$  for RRF4 and RRF5 and from 2,625.8  $g^2$  to 4,857.3  $g^2$  for RRP4 and

RRP5. Permanent environmental variances ranged from 12,136.6  $g^2$  to 32,909.8  $g^2$  for RRF4 and RRF5 and from 9,093.8  $g^2$  to 19,702.0  $g^2$  for RRP4 and RRP5.

RRF4 and RRF5 models presented similar results of correlations between test-day fat yield records on days in milk (Figure 2). Additive genetic and permanent environmental correlations ranged from 0.35 to 1.00 and from 0.22 to 0.97, respectively.

For test-day protein correlations, RRP4 and RRP5 models presented similar ranges from 0.20 to 1.00 for additive genetic and from 0.25 to 0.98 for permanent environmental correlations (Figure 2). High genetic and environmental correlations were found in adjacent test days (close to 1) and decreased with the increasing of distance between test days.

The Spearman rank correlations of EBVs for 305-day fat yield between LMF and RRF4 and between LMP and RRP4 were presented in Table 4. In general, rank correlations for bulls increased from 0.86 to 0.97 with the increasing in classes of bulls' progeny size for fat and from 0.86 to 0.98 for protein (Table 4). Spearman rank correlation between LMF and RRF4, in general, increased from 0.83 to 0.89 for groups of cows with 6 to 10 test days. Rank correlation between LMP and RRP4 increased from 0.81 to 0.86 from 6 to 10 TD.

The average gain in reliability ranged from 3% to 16% for RRF4 and from 6% to 26% for RRP4 for bulls according to the decreasing in the classes of progeny size compared to 305-day lactation models (LMF and LMP) (Table 5). For cows, average gain of reliability was higher than 24% (RRF4) and 38% (RRP4) for

cows compared to LMF and LMP with no differences due to the increasing in the number of test days in lactation.

## **DISCUSSION**

A reduction of the values of  $-2\text{LogL}$ , Akaike information criterion, AIC, and Bayesian information criterion, BIC, as Legendre polynomial order decreases, with a significant change in the log likelihood, indicated random regression models of fourth order Legendre polynomials as the best fit but residual values indicated fifth order as the model (Table 1). AIC, BIC and RV are criteria that tend to indicate models with larger number of parameters (Ducrocq, 2000, López-Romero and Carabaño, 2003; Biassus et al., 2011; Aliloo et al., 2014).

According to Liu et al. (2006) and Aliloo et al. (2014), the choice of the best model is not an easy task because the use of different tests may indicate different models. The authors proposed the use of an index as a tool to indicate the best model. Thus besides AIC, BIC and RV, it was used eigenvalues of genetic and permanent covariance matrices in order to have different methods of comparison to indicate the best model for genetic evaluations. The first three additive eigenvalues explained a sufficiently large proportion of the variances for model RRF4 and RRP4 (Table 2). The highest values of permanent environmental eigenvalues indicated that the permanent environmental effect needed to be modeled with a fifth order Legendre polynomial compared to genetic effect. Pool et al. (2000) used complete lactations for a genetic evaluation of Dutch Holstein cattle to be evaluated by random regression with Legendre polynomial from third to sixth order. The authors reported that a fourth

order for genetic effect and a fifth order for permanent environmental effect allow a simpler covariance function, reducing the number of parameters but as equal orders were suggested in the literature, a fifth order was recommended. Recently Mohamadi and Alijani (2014) conducted a study to compare random regression models with lower orders of Legendre polynomials for genetic and higher for permanent environmental effects. The authors inferred that lower order polynomial for the additive genetic effect than for the permanent environmental effect was better for modeling yield traits in Iranian Holsteins. In Brazil, Araújo et al. (2006) and Biassus et al. (2011) in Minas Gerais State suggested that the best random regression models should be fitted by at least a fourth order Legendre polynomial in order to estimate genetic parameters and breeding values for milk, fat and protein yields. Thus, based on the results, a fourth order of Legendre polynomial for additive genetic and permanent environmental effects for fat and protein yield traits in Brazilian Holsteins is recommended.

Estimates of heritability showed that random regression models presented similar values to the values estimated from 305-day lactation model (LMF and LMP) for 305 day fat and protein yields (Table 3). In literature, the estimates of heritability at 305 days from random regression models in Holstein cattle ranged from 0.29 to 0.41 for fat and from 0.29 to 0.41 for protein yields (Jakobsen *et al.*, 2002; Bohmanova *et al.*, 2008; Biassus et al., 2011; Kheirabadi & Alijani, 2014). In relation to 305-day lactation model, heritability varied from 0.13 to 0.42 for fat yield and from 0.12 to 0.40 for protein yield (Dematawewa & Berger, 1998; Weller & Ezra, 2004; Kim *et al.* 2009).

The heritability of test-days on days in milk for fat estimated from random regression models presented similar values in mid lactation but lower values at the extremes of lactation compared to lactation model (Figure 1). Heritability estimates of protein test-day yields were higher only in the mid lactation (90 to 270 days in milk) than that in the 305-day lactation model. Kim *et al.* (2009) found quite similar results in a population of Holstein in Korea and reported that random regression models presented higher heritability of test days on days in milk than in the 305-day lactation models for fat and protein yields. In previous studies with a Brazilian population of Holstein, Biassus *et al.* (2011) found values ranging from 0.03 to 0.21 and 0.09 to 0.33 for fat and protein yields during days in milk estimated from random regression models. Rzewuska and Strabel (2013) and Abdollahpour *et al.* (2013) reported very close average heritability values from 0.17 to 0.22, 0.14 to 0.23 for fat and protein yields, respectively. Additive genetic and permanent environmental variances of TDF and TDP between DIM 30 and 240 were quite constant for all models and correlated with the highest heritability values in mid lactation (Figure 1). When the extremes of lactation curves were considered, the highest values of permanent environmental correlations were related to the lowest values of additive genetic and heritability. These results were similar to those found by Biassus *et al.* (2011).

The genetic and environmental correlations between test day fat and protein yields (Figure 2) at different days in milk during lactation showed that genetic correlations, in general, were higher than permanent environmental correlations. The figures also indicate that additive genetic correlations for fat



were higher in comparison to protein yield, except between first (DIM 6) and last test days (DIM 270 and 305). Moreover, all genetic and environmental correlation coefficients were above 0.2. Kheirabadi and Alijani (2014) and Biassus *et al.* (2011) reported positive genetic correlation values above 0.34 for fat and 0.36 for protein, which is in agreement with this study.

The Spearman rank correlations of EBVs for 305-day yield between LMF and RRF4 and between LMP and RRP4 of bulls and cow may be observed in two ways (Table 4). First, the rank correlations of EBVs between LMF and RRF4 and between LMP and RRP increased with the increasing in the amount of information of the progeny size of bulls. Second, as bulls' progeny size decreased, there was from 3% to 11% (fat) and from 2% to 11% (protein) of bulls whose EBVs were re-ranked when random regression models were applied. In relation to EBVs for 305 day yields of cows, the increasing in the number of test days from 6 to 10 TD showed an increasing in the correlation.

In terms of reliability of EBVs for 305-day yields, the increasing in the bulls' progeny size was accompanied by decreasing in the average percentage of gain in reliability for fat and protein traits compared to lactation models (LMF and LMP) (Table 5). Bulls with progeny size higher than 100 had an increasing lower than 5% in average but these gains in reliability of EBVs showed the superiority of random regression models even when the amount of information was increased but the most important result was when the bulls' progeny size decreased. Considering bulls with lower progeny size, the percentage was from 8% to 16% for fat and was even higher for protein, between 14% and 26% in average. The percentages of gain of reliability of cows between random

regression and 305-day lactation models were similar for cows with 6, 7, 8, 9 or 10 test days in lactation, which suggested that, in the conditions of present study, the group of cows with less test days presented similar reliability and similar gain in reliability of cows with more than 7 test days in lactation.

## **CONCLUSION**

The random regression model fitted by fourth-order Legendre polynomial is recommended for genetic evaluations of bulls and cows for fat and protein yields of Brazilian Holstein cattle. The main advantages of adoption of random regression model instead of the current 305-day lactation model will be re-ranking of bulls and cows accompanied by the increasing in the reliability of breeding values for 305-day fat and protein yields of Brazilian Holsteins. These advantages are more evident for bulls and cows with less information from test day records.

## **ACKNOWLEDGEMENTS**

We would like to thank the Associação Brasileira de Criadores de Bovinos da Raça Holandesa (ABCBRH) for providing the valuable data. This research was developed with the support of CENAPAD-SP (Centro Nacional de Processamento de Alto Desempenhoem São Paulo), UNICAMP / FINEP project - MCT.

**REFERENCES**

- Abdollahpour R, Moradi-shahrbabak M, Nejati-javaremi A, Vaez-torshizi R, Mrode R. 2013. Genetic analysis of milk yield, fat and protein content in Holstein dairy cows in Iran: Legendre polynomials random regression model applied. *Archiv Tierzucht* **56**, 48.
- Aliloo H, Miraie-Ashtiani SR, Shahrbabak MM, Urioste JI, Sadeghi M. 2014. Accounting for heterogeneity of phenotypic variance in Iranian Holstein test-day milk yield records. *Livestock Science* **167**, 25-32.
- Araújo CV, Torres RA, Costa CV, Filho RAT, Araújo SI, Lopes PS, Regazzi AJ, Pereira CP, Sarmiento JLR. 2006. Random regressions models to describe the genetic variation of milk yield in Holstein breed. *Revista Brasileira de Zootecnia* **35**, 975-981.
- Biassus IO, Cobuci JA, Costa CN, Rorato PRN, Neto JB, Cardoso LL. 2011. Genetic parameters for production traits in primiparous Holstein cows estimated by random regression models. *Revista Brasileira de Zootecnia* **40**, 85-94.
- Brotherstone S, White IMS, Meyer K. 2000. Genetic modeling of daily milk yields using orthogonal polynomials and parametric curves. *Journal of Animal Science* **70**, 407–415.
- Cobuci JA, Costa CN, Neto JB, Freitas AF. 2011. Genetic parameters for milk production by using random regression models with different alternatives of fixed regression modeling. *Revista Brasileira de Zootecnia* **40**, 557-567.
- Costa CN, Melo CMR, Packer IU, Freitas AF, Teixeira NM, Cobuci JA. 2008. Genetic parameters for test day milk yield of first lactation Holstein cows estimated by random regression using Legendre polynomials. *Revista Brasileira*

*de Zootecnia* **37**, 602-608.

Dematawewa CMB, Berger PJ. 1998. Genetic and phenotypic parameters for 305-day yield, fertility, and survival in Holsteins. *Journal of Dairy Science* **81**, 2700–2709.

Ducrocq V. 2000. Calving ease evaluation of French dairy bulls with a heteroskedastic threshold model with direct and maternal effects. *Interbull Bulletin* **25**, 123-130.

Henderson CR, Jr. 1982. Analysis of covariance in the mixed model: Higher-level, nonhomogeneous and random regressions. *Biometrics* **38**, 623–640.

Jakobsen JH, Madsen P, Jensen J, Pedersen J, Christensen LG, Sorensen DA. 2002. Genetic parameters for milk production and persistency for Danish Holstein estimated in random regression models using REML. *Journal of Dairy Science* **85**, 1607-1616.

Jensen J. 2001. Genetic evaluation of dairy cattle using test-day models. *Journal of Dairy Science* **84**, 2803–2812.

Kheirabadi K, Alijani S. 2014. Genetic parameters for milk production and persistency in the Iranian Holstein population by the multitrait random regression model. *Archiv Tierzucht* **57**, 1-12.

Kim B, Lee D, Jeon J, Lee J. 2009. Estimation of genetic parameters for milk production traits using a random regression test-day model in Holstein cows in Korea. *Asian-Australasian Journal of Animal Science* **22**, 923-930.

Kirkpatrick M, HillWG, ThompsonR. 1994. Estimating the covariance structure of traits during growth and aging, illustrated with lactations in dairy cattle. *Genetics Research* **64**, 57-69.

Liu YX, Zhang J, Schaeffer LR, Yang RQ, Zhang WL. 2006. Short communication: optimal random regression models for milk production in dairy cattle. *Journal of Dairy Science* **89**, 2233-2235.

López-Romero P, Carabaño MJ. 2003. Comparing alternative random regression models to analyze first lactation daily milk yield data in Holstein-Friesian cattle. *Livestock Production Science* **82**, 81-96.

Misztal I, Wiggans GR. 1988. Approximation of prediction error variance in large-scale animal models. *Journal of Dairy Science* **71**, 27–32.

Misztal I., Tsuruta S., Aguilar I., Legarra A., Vitezica Z. (2014): Manual for BLUPF90 family of programs. Available: [http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90\\_all1.pdf](http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90_all1.pdf).

Accessed: 30 October, 2014.

Mohammadi A, Alijani S. 2014. Estimation of genetic parameters and comparison of random regression animal and bull models of production traits in the first three lactations of Iranian Holsteins. *Biotechnology in Animal Husbandry* **30**, 261-279.

Muir BL, Kistemaker G, Jamrozik J, Canavesi F. 2007. Genetic parameters for a multiple-trait multiple-lactation random regression test-day model in Italian Holsteins. *Journal of Dairy Science* **90**, 1564-1574.

Pool, M. H., L. L. G. Janss, and T. H. E. Meuwissen. 2000. Genetic parameters of Legendre polynomials for first parity lactation curves. *Journal of Dairy Science* **83**, 2640–2649.

Rzewuska K, Jamrozik J, Żarnecki A, Strabel T. 2011. Genetic parameters of

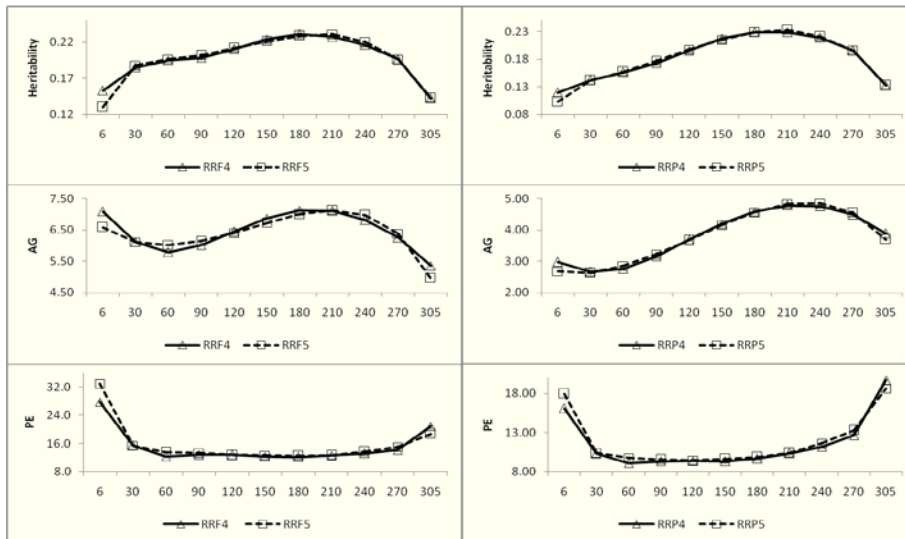
test-day somatic cell scores for the first three lactations of Polish Holstein-Friesian cattle. *Czech Journal of Animal Science* **56**, 381–389.

Schaeffer LR, Dekkers JCM. 1994. Random regressions in animal models for test-day production in dairy cattle. Proc. 5th World Congr. Genet. Appl. Livest. Prod., Guelph, vol. XVIII: 443.

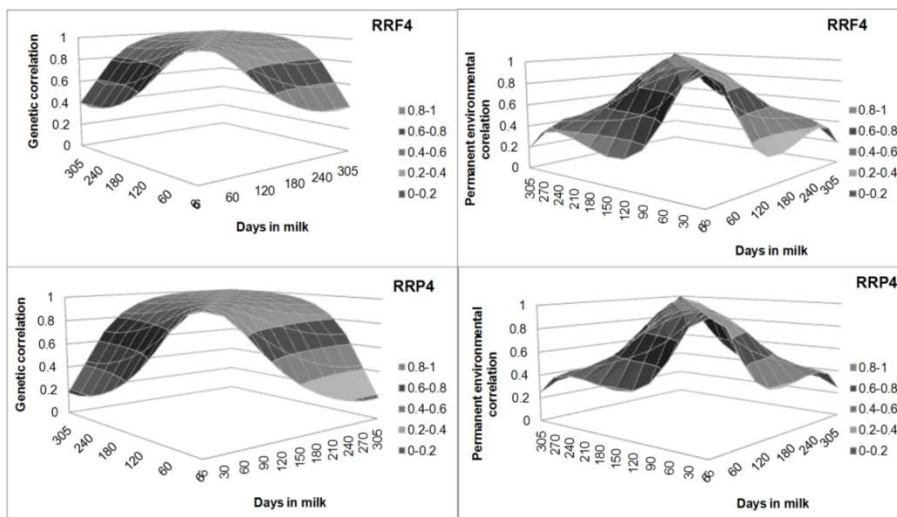
Schaeffer LR, Jamrozik J, Kistemaker GJ, Van Doormaal BJ. 2000. Experience with a test-day model. *Journal of Dairy Science* **83**, 1135-1144.

Weller JI, Ezra E. 2004. Genetic analysis of the Israeli Holstein dairy cattle population for production and non-production traits with a multitrait animal model. *Journal of Dairy Science* **87**, 1519–1527.

Zavadilová L, Wolf J, Štípková M, Němcová E, Jamrozik J. 2011. Genetic parameters for somatic cell score in the first three lactations of Czech Holstein and Fleckvieh breeds using a random regression model. *Czech Journal of Animal Science* **56**, 251–260.



**Figure 1** Heritability, additive genetic (AG) and permanent environmental (PE) variances (x 1,000) on days on milk estimated from random regression models fitted by Legendre polynomials of fourth, RRF4 and RRP4 and fifth, RRF5 and RRP5, orders for fat and protein yields, respectively.



**Figure 2** Genetic correlation (left) and permanent environment correlations (right) between test-days on days in milk estimated from random regression models fitted by Legendre polynomials of fourth, RRF4 and RRP4 and fifth, RRF5 and RRP5, orders for fat and protein yields, respectively.

**Table 1** Number of estimated parameters (p), -2 LogL, Akaike's information criterion (AIC), Bayesian information criterion (BIC), likelihood ratio test (LRT) and chi-square statistics ( $\chi^2$ ) for random regression models

Model	p	-2 Log L	AIC	BIC	RV	LRT	$\chi^2$
RRF4	21	<b>3191455.3</b>	<b>3191497.3</b>	<b>3191569.1</b>	11500	-	-
RRF5	31	3345709.1	3345771.1	3345877.1	11040	154,273.8*	18.30
RRP4	21	<b>3045851.8</b>	<b>3045893.8</b>	<b>3045965.5</b>	5763	-	-
RRP5	31	3199715.1	3199777.1	3199883.0	5469	153,917.5*	18.30

\*  $p < 0.05$ ; RRF4 and RRF5 = random regression model for fat yield fitted by fourth and fifth order Legendre polynomial; RRP4 and RRP5 = random regression model for protein yield fitted by fourth and fifth order Legendre polynomial; RV = residual value; LRT = Likelihood ratio test

**Table 2** Eigenvalues ( $\lambda_i$ ) of the additive genetic (co)variance matrix and the proportion of total variance (%) estimated from random regression models

Trait	Model	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$
Additive genetic variance						
Fat	RRF4	11553.3 (89.5)	1341.5 (10.4)	19.0 (0.15)	1.8 (0.01)	
	RRF5	11561.9 (89.5)	1341.9 (10.4)	19.6 (0.15)	2.3 (0.02)	0.3 (0.002)
Protein	RRP4	6878.2 (88.7)	861.4 (11.1)	15.5 (0.20)	0.9 (0.01)	
	RRP5	6922.0 (89.0)	839.1(10.8)	14.5 (0.19)	1.2 (0.02)	0.2 (0.003)
Permanent environmental variance						
Fat	RRF4	19544.7 (70.8)	4234.9 (15.3)	2748.7 (9.96)	1082.8 (3.9)	
	RRF5	19564.2 (68.8)	4314.8 (15.2)	2769.4 (9.7)	1382.0(4.9)	414.8(1.5)
Protein	RRP4	15353.0 (71.3)	3656.5 (17.0)	1749.0 (8.1)	786.6 (3.7)	
	RRP5	15355.2 (69.6)	3737.5 (17.0)	1726.7 (7.8)	862.9 (3.9)	367.3 (1.7)

RRF4 and RRF5 = random regression model for fat yield fitted by fourth and fifth order Legendre polynomial; RRP4 and RRP5 = random regression model for protein yield fitted by fourth and fifth order Legendre polynomial



**Table 3** Estimates of heritability for 305-day fat and protein yields estimated from 305-day lactation and random regression models

Trait	Models	$h^2$
Fat	LMF	0.24
	RRF4	0.21
	RRF5	0.21
Protein	LMP	0.17
	RRP4	0.20
	RRP5	0.20

LMF = 305-day lactation model for fat yield; LMP = 305-day lactation model for protein yield; RRF4 and RRF5 = random regression model for fat yield fitted by fourth and fifth order Legendre polynomial; RRP4 and RRP5 = random regression model for protein yield fitted by fourth and fifth order Legendre polynomial

**Table 4** Spearman rank correlations of EBVs for 305-day fat yield between LMF and RRF4 and between LMP and RRP4 of bulls and cows according to classes of progeny size and number of test days in lactation

Bulls		Models	
Progeny size	Number of bulls	RRF4	RRP4
200 – 399	29	0.97	0.98
100 – 199	74	0.96	0.94
50 – 99	154	0.95	0.92
25 – 49	175	0.90	0.87
10 – 24	352	0.86	0.86
Cows		Models	
Number of test days	Number of cows	RRF4	RRP4
10	9449	0.89	0.86
9	9569	0.88	0.86
8	5433	0.85	0.85
7	3534	0.85	0.85
6	2243	0.80	0.81

LMF = 304-day lactation model for 305-day fat yield; RRF4 and RRP4 = random regression model for fat and protein yield fitted by fourth order Legendre polynomial

**Table 5** Average percentage of increasing in reliability (range in parenthesis) of EBVs for 305-day yield estimated from random regression model compared to 305-day lactation model

Bulls		Models			
Progeny size	Number of bulls	LMF <sup>1</sup>	RRF4 <sup>2</sup>	LMP <sup>1</sup>	RRP4 <sup>2</sup>
200 – 399	29	0.90 ± 0.06	+3% (1-10%)	0.87 ± 0.07	+6% (2-18%)
100 – 199	74	0.85 ± 0.05	+5% (2-13%)	0.81 ± 0.06	+8% (3-22%)
50 – 99	154	0.73 ± 0.10	+8% (3-24%)	0.68 ± 0.11	+14% (5-39%)
25 – 49	175	0.62 ± 0.10	+12% (3-27%)	0.56 ± 0.10	+19% (5-43%)
10 – 24	352	0.44 ± 0.12	+16% (7-39%)	0.38 ± 0.12	+26% (10-63%)
Cows		Models			
Number of test-days	Number of cows	LMF <sup>1</sup>	RRF4 <sup>2</sup>	LMP <sup>1</sup>	RRP4 <sup>2</sup>
10	9449	0.37 + 0.08	+26% (0-427%)	0.31 + 0.09	+38% (0-215%)
9	9569	0.36 + 0.08	+26% (12-85%)	0.31 + 0.09	+39% (17-120%)
8	5433	0.34 + 0.08	+27% (6-159%)	0.29 + 0.09	+40% (10-214%)
7	3534	0.34 + 0.08	+26% (10-75%)	0.29 + 0.09	+39% (15-111%)
6	2243	0.33 + 0.09	+24% (9-70%)	0.28 + 0.09	+38% (14-105%)

<sup>1</sup> average and standard-deviation of reliability values of EBVs from lactation models; <sup>2</sup>

average percentage of proportional gains of reliability compared to lactation model with range in (parenthesis); LMF = 305-day lactation model for 305-day fat yield; LMP = 305-day lactation model for 305-day protein yield; RRF4 and RRP4 = random regression models fitted by fourth-order Legendre polynomial for fat and protein yield, respectively



## 5. CAPÍTULO V



## CONSIDERAÇÕES FINAIS

As diferentes estruturas de dados, formadas ao considerar apenas vacas com 4, 6, 8 ou 10 registros de produção de leite na lactação, influenciaram os resultados estimados pelos modelos de regressão aleatória. O aumento na ordem dos polinômios de terceira para quinta ordem não alterou a estimação dos valores genéticos aos 305 dias ou a confiabilidade dos mesmos. No entanto, a diferenciação das estruturas dos dados diminuiu a confiabilidade dos valores genéticos de um grupo de touros à medida que as restrições aumentaram de vacas com 4 registros para vacas com 10 registros na lactação. Esses resultados indicam que o aumento nas restrições dos animais modifica a estrutura dos dados a ponto de diminuir muito o número de animais avaliados e, conseqüentemente, a confiabilidade dos valores genéticos. A estrutura com pelo menos 10 registros por vaca na lactação apresentou a menor confiabilidade média e também representou menos de  $\frac{1}{4}$  do tamanho da base de dados comparado com a base que continha pelo menos 4 registros.

Ao comparar modelos de regressão aleatória de quarta e quinta ordens de polinômios de Legendre com modelos de lactação aos 305 dias, foi possível perceber que a distribuição dos valores genéticos acumulados aos 305 dias pelo modelo de regressão aleatória foi diferente dos valores estimados pelo modelo de lactação. A correlação de posto dos valores genéticos de touros estimado pelo modelo de lactação e pelos modelos de regressão aumentou conforme diminuiu o número de filhas por touro, assim como diminuiu paravacas, conforme o número de controles na lactação foi menor. Isso indica um re-ordenamento de touros pelos modelos de regressão aleatória, principalmente para touros com menos informação de progênie. Além disso, as confiabilidades dos valores genéticos estimados pelos modelos de regressão aleatória foram maiores do que aqueles estimados pelos modelos de lactação, e, principalmente para touros com menos filhas avaliadas, ou seja, touros com menos informação. Isso representa uma grande vantagem num programa de melhoramento genético, principalmente se os touros com menos informação são touros jovens que poderiam ser avaliados mais cedo, aumentando o ganho genético anual pela diminuição do intervalo de gerações. Em relação às vacas com números diferentes de controles, foi possível observar que vacas com 6, 7, 8, 9 ou 10 registros no dia do controle apresentaram a mesma média de ganho de confiabilidade comparado com o modelo de lactação aos 305 dias. Esse resultado sugere a possibilidade de estimar valores genéticos sem a necessidade de aumentar o número de controles na lactação.



## 6. REFERÊNCIAS BIBLIOGRÁFICAS

ABDULLAHPOUR, R. et al. Genetic analysis of milk yield, fat and protein content in Holstein dairy cows in Iran: Legendre polynomials random regression model applied. **Archiv Tierzucht**, Dummerstorf, v. 56 n. 48, p. 497-508, 2013.

AIKMAN, P. C.; REYNOLDS, C. K.; BEEVER, D. E. Diet Digestibility, Rate of passage, and eating and rumination behavior of Jersey and Holstein Cows. **Journal of Dairy Science**, Champaign, v. 91, n. 3, p. 1103-1114, 2008.

AKAIKE, H. Information theory and an extension of the maximum likelihood principle. In: INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY, 2., 1973, Budapest, Hungary. **Proceedings...** Budapest: Akademia Kiado, 1973. p. 267-281.

ALI, T. E.; SCHAEFFER, L. R. Accounting for covariances among test day milk yields in dairy cows. **Canadian Journal of Animal Science**, Ottawa, v. 67, n. 3, p. 637-644, 1987.

ALILOO, H. et al. Accounting for heterogeneity of phenotypic variance in Iranian Holstein test-day milk yield records. **Livestock Science**, Amsterdam, v. 167, p. 25-32, 2014.

ANUALPEC. **Anuário da Pecuária Brasileira**. São Paulo: FNP, 2011. 385 p.

ARAÚJO, C. V. et al. Uso de modelos de regressão aleatória para descrever a variação genética da produção de leite na raça Holandesa. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 35, n. 3, p. 975-981, 2006.

ARDALAN, M.; DEHGHAN-BANADAKY, M.; REZAYAZDI, K. Milk yield persistency and its relationship with health problems in Holstein dairy cows supplemented with different levels of ruminally protected methionine and choline. **Archiv Tierzucht**, Dummerstorf, v. 53 n. 3, p. 266-276, 2010.

BAÑOS, G. et al. Genetic relationship between first-lactation body energy and later-life udder health in dairy cattle. **Journal of Dairy Science**, Champaign, v. 89, n. 6, p. 2222–2232, 2006.



BIASSUS, I. O. et al. Genetic parameters for production traits in primiparous Holstein cows estimated by random regression models. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 40, n. 1, p. 85-94, 2011.

BIASSUS, I. O. et al. Persistence in milk, fat and protein production of primiparous Holstein cows by random regression models. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 39, n. 12, p. 2617-2624, 2010.

BIGNARDI, A. B. et al. Random regression models to estimate test-day milk yield genetic parameters Holstein cows in Southeastern Brazil. **Livestock Science**, Amsterdam, v. 123, n. 1, p. 1-7, 2009.

BIGNARDI, A. B. et al. Random regression models using different functions to model test-day milk yield of Brazilian Holstein cows. **Genetics and Molecular Research**, Ribeirão Preto, v. 10, n. 4, p. 3565-3575, 2011.

BILAL, G.; KHAN, M. S. Use of test-day milk yield for genetic evaluation in dairy cattle: a review. **Pakistan Veterinary Journal**, Faisalabad, v. 29, n. 1 p. 35-41, 2009.

BOHMANOVA, J. et al. Comparison of random regression models with Legendre polynomials and linear splines for production traits and somatic cell score of Canadian Holstein cows. **Journal of Dairy Science**, Champaign, v. 91, n. 9, p. 3627-3638, 2008.

BORMANN, J. et al. Within-herd effects of age at test-day and lactation stage on test-day yields. **Journal of Dairy Science**, Champaign, v. 86, n. 11, p. 3765-3774, 2003.

BOZDOGAN, H. Akaike's information criterion and recent developments in information complexity. **Journal of Mathematical Psychology**, New York, v. 44, n. 1, p. 62-91, 2000.

BRASIL. Ministério da Agricultura. Secretaria Nacional de Produção Agropecuária. Portaria nº 45, de 10 de outubro de 1986. Normas técnicas para execução do serviço de controle leiteiro em bovídeos. **Diário Oficial [da] República Federativa do Brasil**. Brasília, v. 124, n. 198, p. 15532-15535, out. 1986. Seção 1.

BROTHERSTONE, S.; WHITE, I. M. S.; MEYER, K. Genetic modelling of daily yield using orthogonal and parametric curves. **Animal Science**, v. 70, n. 3, p. 407-415, 2000.

ÇANKAYA, S. et al. Comparison of some random regression models for first lactation test day milk yields in Jersey cows and estimating of genetic parameters. **Kafkas Üniversitesi Veteriner Fakültesi Dergisi**, Kars, v. 20, n. 1, p. 5-10, 2014.

ÇİLEK, S. Estimation of adjustment factors for standardizing lactations to mature age and 305 day of milk yield of Holstein cattle reared at Polatli State farm in Turkey. **Journal of Animal and Veterinary Advances**, Faisalabad, v. 7, n. 9, p. 1056-1060, 2008.

COBUCI, J. A.; COSTA, C. N. Persistency of lactation using random regression models and different fixed regression modeling approaches. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 41, n. 9, p. 1996-2004, 2012.

COBUCI, 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**, Viçosa, MG, v. 40, n. 3, p. 557-567, 2011.

COBUCI, J. A. et al. Análises da persistência na lactação de vacas da raça Holandesa, usando produção no dia do controle e modelo de regressão aleatória. **Brazilian Journal of Animal Science**, Viçosa, v. 33, n. 3, p. 546-554, 2004.

COBUCI, J. A. et al. Genetic evaluation for persistency of lactation in Holstein cows using a random regression model. **Genetics and Molecular Biology**, São Paulo, v. 30, n. 2, p. 349-355, 2007.

COBUCI, J. A. et al. Persistência na lactação - uma revisão. **Archivos Latinoamericanos de Producción Animal**, Mayaguez, v. 11, n. 3, p. 163-173, 2003.

COSTA, C. N. et al.. **Sumário Nacional de Touros da Raça Holandesa – 2012**. Juiz de Fora: Embrapa Gado de Leite, 2012. 40 p. (Documentos, 158).

COSTA, C. N. 2005. **Reflexões sobre a gestão zootécnica**: o controle leiteiro na raça Holandesa. Alguma lição? MilkPoint. Disponível em: <http://www.milkpoint.com.br/radar-tecnico/melhoramento-genetico/reflexoes-sobre-a-gestao-zootecnica-o-controle-leiteiro-na-raca-holandesa-alguma-licao-25542n.aspx>. Acesso em: 24 nov. 2014.

COSTA, C. N. et al. Estimation of genetic parameters of test day fat and protein yields in Brazilian Holstein cattle using an autoregressive multiple lactation animal model. **South African Journal of Animal Science**, Pretoria, v. 39, n. 5, p. 165-168, 2009. (Supplement 1).

COSTA, C. N. et al. Genetic parameters for test day milk yield of first lactation Holstein cows estimated by random regression using Legendre polynomials. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 37, n. 4, p. 602-608, 2008.

DE ROOS, A. P. W.; HARBERS, A. G. F.; DE JONG, G. Random regression test-day model in The Netherlands. **Interbull Bulletin**, Uppsala, Sweden, v. 27, p. 155-158, 2001.

DEKKERS, J. C. M.; TEM HAG, J. H.; WEERSINK, A. Economic aspects of persistence of lactation in dairy cattle. **Livestock Production Science**, Amsterdam, v. 53, n. 3, p. 237-252, 1998.

DETILLEUX, J. C. Genetic factors affecting susceptibility to udder pathogens. **Veterinary Microbiology**, Amsterdam, v. 134, n. 1-2, p. 157-164, 2009.

DOBSON, H. et al. The High-producing Dairy Cow and its Reproductive Performance. **Reproduction in Domestic Animals**, Berlin, v. 42, p. 17-23, 2007.

DORNELES, C. K. P. et al. Estimation of genetic parameters for Holstein cows milk production by random regression. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, Belo Horizonte, v. 61, n. 2, p. 407-412, 2009.

DZOMBA, E. F. et al. Random regression test-day model for the analysis of dairy cattle production data in South Africa: creating the framework. **South African Journal of Animal Science**, Pretoria, v. 40, n. 4, p. 273-284, 2010.

EVERETT, R. W.; CARTER, H. W. Accuracy of test interval method of calculating dairy herd improvement association records. **Journal of Dairy Science**, Champaign, v. 51, n. 12, p. 1936-1941, 1968.

FAO. **Base de dados estatísticos da Food and Agriculture Organization of the United Nations**. Roma: FAOSTAT. Disponível em: <[www.faostat.fao.org](http://www.faostat.fao.org)>. Acesso em: 14 dez. 2014.

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

FERREIRA, W. J. et al. Utilização da produção de leite no dia do controle na avaliação genética em gado de leite - uma revisão. **Archivos Latinoamericanos de Producción Animal**, Mayaguez, v. 10, n. 1, p. 46-53, 2002.

HAMED, M. K. Accuracy of bimonthly and trimonthly milk recording systems for dairy cattle in Egypt. **Annals of Agricultural Science, Moshtohor Journal**, Benha, Egypt, v. 33, p. 669-679, 1995.

HENDERSON, C. R. Jr. Analysis of covariance in the mixed model: Higher-level, nonhomogeneous and random regressions. **Biometrics**, Alexandria Va, v. 38, n. 3, p. 623-640, 1982.

HOLANDA, M. C. R. et al. Modelo autorregressivo para avaliação da produção leiteira em rebanho Holandês. **Revista Brasileira de Saúde e Produção Animal**, Salvador, v. 12, n. 3, p. 618-629, 2011.

HUELSENBECK, J. P.; BULL, J. J. A likelihood ratio test to detect conflicting phylogenetic signal. **Systematic Biology**, Washington, v. 45, n. 1, p. 92-98, 1996.

JAKOBSEN, J. H. et al. Genetic parameters for milk production and persistency for Danish Holstein estimated in random regression models using REML. **Journal of Dairy Science**, Champaign, v. 85, n. 6, p. 1607-1616, 2002.

JAMROZIK, J.; BOHMANOVA, J.; SCHAEFFER, L. R. Selection of locations of knots for linear splines in random regression test-day models. **Journal of Animal Breeding and Genetics**, Malden, MA, USA, v. 127, n. 2, p. 87–92, 2010.

JAMROZIK, J.; SCHAEFFER, L. R. Estimates of genetic parameters for a test day model with random regressions for yield traits of first lactation Holsteins. **Journal of Dairy Science**, Champaign, v. 80, n. 4, p. 762-770, 1997.

JAMROZIK, J.; SCHAEFFER, L. R.; DEKKERS, J. C. M. Genetic evaluation of dairy cattle using test day yields and random regression model. **Journal of Dairy Science**, Champaign, v. 80, n. 6, p. 1217–1226, 1997a.

JAMROZIK, J. et al. Multiple trait random regression test day model for production traits. **Interbull Bulletin**, Uppsala, Sweden, v. 16, p. 43-47, 1997b.

MORAIS JÚNIOR, N. N. et al. Avaliação dos fatores oficiais de correção da produção até o primeiro controle leiteiro em vacas Holandesas. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 31, n. 2, p. 924-933, 2002 (suplemento).

KAYA, I.; AKBAS, Y.; UZMAY, C. Estimation of breeding values for dairy cattle using test-day milk yields. **Turkish Journal of Veterinary and Animal Sciences**, [Turkey], v. 27, p. 459-464, 2003.

KHORSHIDIE, R. et al. Genetic trends for 305-day milk yield and persistency in Iranian Holsteins. **Livestock Science**, Amsterdam, v. 144, n. 3, p. 211–217, 2012.

KIM, B. et al. Estimation of genetic parameters for milk production traits using a random regression test-day model in Holstein cows in Korea. **Asian-Australasian Journal of Animal Science**, Seoul, v. 22, n. 7, p. 923-930, 2009.

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

KISTEMAKER, G. J. The Canadian test-day model using Legendre polynomials. **Interbull Bulletin**, Uppsala, Sweden, v. 31, p. 202-204, 2003.

LIU, Y. X. et al. Short communication: optimal random regression models for milk production in dairy cattle. **Journal of Dairy Science**, Champaign, v. 89, n. 6, p. 2233–2235, 2006.

LIU, Z. et al. Application of random regression model to genetic evaluations of test day yields and somatic cell scores in dairy cattle. **Interbull Bulletin**, Uppsala, Sweden, v. 27, p. 159–166, 2001.

MARTINS, A. M. et al. Somatic cell score genetic parameter estimates of dairy cattle in Portugal using fractional polynomials. **Journal of Animal Science**, Champaign, v. 89, n. 5, p. 1281-1285, 2011.

MELO, C. M. R. et al. Parâmetros genéticos para as produções de leite no dia do controle e da primeira lactação de vacas da raça Holandesa. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 34, n. 3, p. 796-806, 2005.

MELO, C. M. R. et al. Valores genéticos para as produções de leite do dia do controle e da lactação na raça Holandesa com diferentes modelos estatísticos. **Revista Brasileira de Zootecnia**, Viçosa, MG, v. 36, n. 5, p. 1295-1303, 2007.

MEYER, K. Estimating covariance functions for longitudinal data using a random regression model. **Genetic Selection Evolution**, Paris, v. 30, p. 221-240, 1998.

MOHAMMADI, A.; ALIJANI, S. Estimation of genetic parameters and comparison of random regression animal and bull models of production traits in the first three lactations of Iranian Holsteins. **Biotechnology in Animal Husbandry**, Zemun, Sérvia, v. 30, n. 2, p. 261-279, 2014.

MOSTERT, B. E. et al. Fixed regression test-day models for South African dairy cattle for participation in international evaluations. **South African Journal of Animal Science**, Pretoria, v. 36, p. 58-70, 2006b.

MRODE, R.; CARVALHEIRA, J.; SCHAEFFER, L. Test day models. In: GROSU, H.; LUNGU, S.; OLTENACU, P. A. **History of genetic evaluation methods in dairy cattle**. Bucharest, Romania: [s.n.], 2013.

MRODE, R. et al. Joint estimation of genetic parameters for test-day somatic cell count and mastitis in the United Kingdom. **Journal of Dairy Science**, Champaign, v. 95, n. 08, p. 4618-4628, 2012.

MRODE, R. A. **Linear models for the prediction of animal breeding values**. 2. ed. Wallingford: CABI Publishing, 2005. 344 p.

MRODE, R. A.; SWANSON, G. J. T.; PAGET, M. F. Implementation of the test day model for production traits in the UK. **Interbull Bulletin**, Uppsala, Sweden, v. 31, p. 193–196, 2003.

MUIR, B. L. et al. Genetic parameters for a multiple-trait multiple-lactation random regression test-day model in Italian Holsteins. **Journal of Dairy Science**, Champaign, v. 90, n. 3, p. 1564-1574, 2007.

NARANCHULUUM, G. et al. Selecting the desirable method for predicting 305-day lactation yields in Mongolia. **Animal Science Journal**, Richmond, v. 82, n. 3, p. 383-389, 2011.

NUNEZ-ANTON, V.; ZIMMERMAN, D. L. Modeling nonstationary longitudinal data. **Biometrics**, Washington, v. 56, n. 3, p. 699-705, 2000.

PEREIRA, R. J. et al. Funções de covariância para produção de leite no dia do controle em bovinos Gir leiteiro. **Pesquisa Agropecuária Brasileira**, Brasília, v. 45, n. 11, p. 1303-1311, 2010.

POOL, M. H.; JANSS, L. L. G.; MEUWISSEN, T. H. E. Genetic parameters of Legendre polynomials for first parity lactation curves. **Journal of Dairy Science**, Champaign, v. 83, n. 11, p. 2640-2649, 2000.

POOL, M. H.; MEUWISSEN, T. H. E. Effects of Random Regression Test-Day Models on EBVs and Genetic Trends in Persistency. **Interbull bulletin**, Uppsala, Sweden, v. 27, p. 184-188, 2001.

PTAK, E.; SCHAEFFER, L. R. Use of test day yields for genetic evaluation of dairy bulls and cows. **Livestock Production Science**, Amsterdam, v. 34, n. 1-2, p. 23-34, 1993.

REENTS, R. et al. Impact of application of a test-day model to dairy production traits on genetic evaluation of cows. **Interbull Bulletin**, Uppsala, Sweden, v. 17, p. 49-54, 1998.

SAWALHA, R. M. et al. Genetic evaluation of dairy cattle with test-day models with autoregressive covariance structures and with a 305-day model. **Journal of Dairy Science**, Champaign, v. 88, n. 9, p. 3346-3353, 2005.

SCHAEFFER, L. R. Application of random regression models in animal breeding. **Livestock Production Science**, Amsterdam, v. 86, n. 1-3, p. 35-45, 2004.

SCHAEFFER, L. R.; DEKKERS, J. C. M. Random regressions in animal models for test-day production in dairy cattle. In: WORLD CONGRESS OF GENETICS APPLIED LIVESTOCK PRODUCTION, 5., 1994, Guelph. **Proceedings...** Guelph: University of New Guelph, 1994. p. 443-44. v. 18.

SCHWARZ, G. Estimating the dimensional of a model. **Annals of Statistics**, Hayward, v. 6, n. 2, p. 461-464, 1978.

STRABEL, T.; PTAK, E.; SZYDA, J. et al. Multiple-lactation random regression test-day model for Polish Black and White cattle. **Interbull Bulletin**, Uppsala, Sweden, n. 32, p. 133-136, 2004.

STRANDÉN, I.; MÄNTYSAARI, E. A. Comparison of two methods to approximate reliabilities of breeding values from test day model. In: MEETING OF THE EUROPEAN ASSOCIATION FOR ANIMAL PRODUCTION, 55., 2004, Bled, Slovenia. **Proceedings...** Bled, Slovenia: [s.n.], 2004.

SWALVE, H. H. Theoretical basis and computational methods for different test-day genetic evaluation models. **Journal of Dairy Science**, Champaign, v. 83, n. 5, p. 1115-1124, 2000.

TAKMA, Ç.; AKBAŞ, Y. Comparison of Fitting Performance of Random Regression Models to Test Day Milk Yields in Holstein Friesians. **Kafkas Universitesi Veteriner Fakültesi Dergisi**, Turkey, v. 15, n. 2, p. 261-266, 2009.



TAKMA, Ç.; AKBAŞ, Y. Estimates of genetic parameters for test day milk yields of a Holstein Friesian herd in Turkey with random regression models. **Archiv Tierzucht**, Dummerstorf, v. 50, n. 4, p. 327-336, 2007.

TEKERLI, M. et al. Factors affecting the shape of lactation curves of Holstein cows from the Balıkesir Province of Turkey. **Journal of Dairy Science**, Champaign, v. 83, n. 6, p. 1381-1386, 2000.

TOGASHI, K.; LIN, C. Y.; YOKOUCHI, K. Overview of genetic evaluation in dairy cattle. **Animal Science Journal**, Tokyo, v. 75, n. 4, p. 275–284, 2004.

TORSHIZI, E. M. et al. Evaluation of test day milk yield in Iranian primiparous Holstein using different random regression models. **Journal of Animal Science Advances**, [S.l.], v. 2, n. 8, p. 664-677, 2012.

TORSHIZI, M. E. et al. Analysis of Test Day Milk Yield by Random Regression Models and Evaluation of Persistency in Iranian Dairy Cows. **Iranian Journal of Applied Animal Science**, Rasht, Irã, v. 3, n. 1, p. 67-76, 2013.

WASIKE, C. B.; KAHI, A. K.; PETERS, K. J. Modelling of lactation curves of dairy cows based on monthly test day milk yield records under inconsistent milk recording scenarios. **Animal**, Cambridge, v. 5, n. 11, p. 1780-1790, 2011.

WERF, J. H. J.; GODDARD M. E.; MEYER, K. The use of covariance functions and random regressions for genetic evaluation of milk production based on test day records. **Journal of Dairy Science**, Champaign, v. 81, p. 3300-3308, 1998.

WHITE, I. M.; THOMPSON, R.; BROTHERSTONE, S. Genetic and environmental smoothing of lactation curves with cubic splines. **Journal of Dairy Science**, Champaign, v. 82, n. 3, p. 632-638, 1999.

WIGANS, G. R. Estimating daily yields of cows milked three times a day. **Journal of Dairy Science**, Champaign, v. 69, n. 11, p. 2935-2940, 1986.

WILMINK, J. B. M. Efficiency of selection for different cumulative milk, fat and protein yields in first lactation. **Livestock Production Science**, Amsterdam, v. 17, n. 3, p. 211-224, 1987.

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

YAMAZAKI, T. et al. Genetic correlations between milk production traits and somatic cell scores on test day within and across first and second lactations in Holstein cows. **Livestock Science**, Amsterdam, v. 152, n. 2-3, p. 120-126, 2013.

ZADRA, L. F. Persistência da lactação de bovinos leiteiros. **Pesquisa & Tecnologia**, Campinas, v. 9, n. 2, 2012.

ZAVADILOVÁ L.; JAMROZIK J.; SCHAEFFER L. R. Genetic parameters for test-day model with random regressions for production traits of Czech Holstein cattle. **Czech Journal of Animal Science**, Praha, v. 50, n. 4, p. 142–154, 2005.

ZAVADILOVÁ, L. et al. Genetic parameters for somatic cell score in the first three lactations of Czech Holstein and Fleckvieh breeds using a random regression model. **Czech Journal of Animal Science**, Praha, v. 56, n. 6, p. 251–260, 2011.

## 7. VITA

Alessandro Haiduck Padilha nasceu em 26 de abril de 1981 no município de Erechim no Estado do Rio Grande do Sul. É filho de Luis Carlos Nunes Padilha e Carmen Haiduck Padilha.

Realizou o ensino fundamental na Escola Municipal Elena Canho Sampaio de Novo Hamburgo e concluiu o segundo grau com Habilitação Tradutor e Intérprete de Inglês no Colégio 25 de Julho também em Novo Hamburgo.

Em 2002 ingressou no curso de Agronomia na Universidade Federal do Rio Grande do Sul (UFRGS), onde no ano de 2009 obteve o diploma de graduação.

Em março de 2009 iniciou o curso de mestrado no Programa de Pós-Graduação em Zootecnia da Faculdade de Agronomia da UFRGS, o qual foi concluído em 2010.