

# Reliability of breeding values between random regression and 305-day lactation models

Alessandro Haiduck Padilha<sup>(1)</sup>, Jaime Araujo Cobuci<sup>(1)</sup>, Darlene dos Santos Daltra<sup>(1)</sup> and José Braccini Neto<sup>(1)</sup>

<sup>(1)</sup>Universidade Federal do Rio Grande do Sul, Faculdade de Agronomia, Departamento de Zootecnia, Avenida Bento Gonçalves, nº 7.712, CEP 91540-000 Porto Alegre, RS, Brazil. E-mail: ahaiduck@yahoo.com.br, jaime.cobuci@ufrgs.br, jose.braccini@ufrgs.br, darlenedaltro@gmail.com

**Abstract** – The objective of this work was to verify the gain in reliability of estimated breeding values (EBVs), when random regression models are applied instead of conventional 305-day lactation models, using fat and protein yield records of Brazilian Holstein cattle for future genetic evaluations. Data set contained 262,426 test-day fat and protein yield records, and 30,228 fat and protein lactation records at 305 days from first lactation. Single trait random regression models using Legendre polynomials and single trait lactation models were applied. Heritability for 305-day yield from lactation models was 0.24 (fat) and 0.17 (protein), and from random regression models was 0.20 (fat) and 0.21 (protein). Spearman correlations of EBVs, between lactation models and random regression models, for 305-day yield, ranged from 0.86 to 0.97 and 0.86 to 0.98 (bulls), and from 0.80 to 0.89 and 0.81 to 0.86 (cows), for fat and protein, respectively. Average increase in reliability of EBVs for 305-day yield of bulls ranged from 2 to 16% (fat) and from 4 to 26% (protein), and average reliability of cows ranged from 24 to 38% (fat and protein), which is higher than in the lactation models. Random regression models using Legendre polynomials will improve genetic evaluations of Brazilian Holstein cattle due to the reliability increase of EBVs, in comparison with 305-day lactation models.

**Index terms:** breeding value, correlation, Legendre polynomials, reliability.

## Confiabilidade de valores genéticos entre modelos de regressão aleatória e de lactação aos 305 dias

**Resumo** – O objetivo deste trabalho foi verificar o ganho em confiabilidade dos valores genéticos estimados (VGE), quando os modelos de regressão aleatória são aplicados em lugar dos modelos convencionais de lactação até os 305 dias, tendo-se utilizado registros de produção de gordura e proteína de bovinos da raça Holandesa no Brasil, para futuras avaliações genéticas. O conjunto de dados continha 262.426 registros de produção de gordura e proteína no dia do controle e 30.228 registros de produção acumulada de gordura e proteína até os 305 dias de lactação. Aplicaram-se modelos de regressão aleatória unicaracterística, com os polinômios de Legendre, e modelos unicaracterística de lactação aos 305 dias. A herdabilidade estimada para as produções até os 305 dias pelos modelos de lactação foi 0,24 (gordura) e 0,17 (proteína), e, pelos modelos de regressão aleatória, foi 0,20 (gordura) e 0,21 (proteína). Correlações de Spearman dos VGEs entre os modelos de lactação e de regressão aleatória, para a produção aos 305 dias, variaram de 0,86 a 0,97 e 0,86 a 0,98 (touro) e de 0,80 a 0,89 e 0,81 a 0,86 (vacas), quanto à gordura e à proteína, respectivamente. O aumento médio em confiabilidade dos VGEs para a produção aos 305 dias variou de 2 a 16% (gordura) e de 4 a 26% (proteína), em touros, e variou de 24 a 38% (gordura e proteína) em vacas, o que é superior àquela obtida pelos modelos de lactação. Os modelos de regressão aleatória com uso dos polinômios de Legendre melhorarão a avaliação genética de bovinos da raça Holandesa no Brasil, em razão do aumento de confiabilidade dos VGEs, em comparação aos modelos de lactação de 305 dias.

**Termos para indexação:** valor genético, correlação, polinômios de Legendre, confiabilidade.

### Introduction

The breeding objectives for Holstein cattle are defined by a list of economically relevant traits, namely milk volume, fat, and protein yields, for instance (Banga et al., 2014). In Brazil, genetic evaluations of

Holstein cattle for these yields have been carried out using a 305-day lactation model (Ferreira et al., 2003; Costa et al., 2009; Biassus et al., 2011). Alternatively, many other approaches, as repeatability, autoregressive or random regression models (Melo et al., 2007; Costa et al., 2008, 2009; Bignardi et al., 2011) have been

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 based on test-day models are that they permit to model the shape of the lactation curve (Schaeffer et al., 2000), besides accounting for environmental factors that affect test-day records of cows at different stages of lactation with more accuracy (Jensen, 2001). Among test-day models, the random regression models have been proposed for genetic evaluations in the literature (Kirkpatrick et al., 1994; DeGroot et al., 2007; Bignardi et al., 2009). In fact, 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). Costa et al. (2008), Biassus et al. (2011), and Cobuci et al. (2011) 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 to substitute the current 305-day lactation model. However, there are few studies on milk and its components, as fat and protein, in tropical countries as Brazil, using random regression models in the genetic evaluations of Holstein breed (Costa et al., 2008; Biassus et al., 2010). In general, the random regression models using Legendre polynomials of fourth, fifth and sixth orders have been indicated as a good option for conducting genetic evaluations of Holstein cattle; however, there are no comparisons between random regression and lactation models as for the advantages in the gain of reliability of breeding values from studies on Brazilian Holsteins (Costa et al., 2008; Biassus et al., 2010, 2011).

The objective of this work was to verify the gain in reliability of estimated breeding values (EBVs), when random regression models are applied instead of conventional 305-day lactation models, using fat and protein yield records of Brazilian Holstein cattle to be used in future genetic evaluations.

## Materials and Methods

Data consisted of fat and protein milk yield collected by the technicians of the milk control and genealogy service of the Associação Brasileira de Criadores de Bovinos da Raça Holandesa (ABCBRH Brazilian

association of Holstein breeders), and its state affiliates between 1990 and 2011. The data set comprised test-day milk yield and 305-day milk yield records. At first, pedigree data was checked for inconsistencies. A minimum of six test-day records, obtained between six and 305 days in milk, were edited for cows aged 18 to 48 months during the first lactation. Abnormal yield values or outliers were checked by graphical techniques as normal probability plots and boxplots, as well as by median, mean, mode, skewness and kurtosis values. Test-day records were removed, if fat and protein yields were out of the range from 258.4 to 1,510 g, and from 312.0 to 1,314.8 g, respectively. The records of 305-day fat and protein yields were deleted, if they were out of the range from 102 to 392 kg, and from 106 to 349 kg, 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. Contemporary groups of herd-year-season of calving (305-day yield records) and herd-year-month of test (test-day records) which did not have at least four records and progeny of bulls, with at least two daughters in two different herds, were eliminated.

After editions in dataset, 262,426 test-day fat and test-day protein yield records were used to apply the random regression models, and 30,228 lactation records were used to apply the 305-day lactation models. Milk traits considered in the present analyses were 305-day fat yield (305F) and 305-day protein yield (305P).

Test-day fat and test-day protein yield records were used in single trait random regression animal model, named as RRF4 and RRF5 for fat yield, and RRP4 and RRP5 for protein yield, fitted by fourth and fifth orders of Legendre polynomials, respectively. The random regression model used to estimate genetic parameters and EBVs for 305F and 305P was as follows:

$$y_{ijklm} = HYM_i + \sum_{k=1}^n \varphi_k(d_t) \beta_{mk} + \sum_{k=1}^n \varphi_k(d_t) u_{jk} + \sum_{k=1}^n \varphi_k(d_t) pe_{jk} + \varepsilon_{ijklm}$$

in which:  $y_{ijkl}$  is the  $l^{\text{th}}$  test-day record (fat or protein) of cow  $j$ , made on day  $t$ , within HYM (herd-year-month of test) subclass  $i$ ;  $\beta_{mk}$  is the  $k^{\text{th}}$  fixed regression coefficient specific for the  $m^{\text{th}}$  subclass of calving age-season

classes;  $\text{HYM}_i$  is the fixed effect of 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  $j$ ;  $\phi_k(d_t)$  are the Legendre polynomials for the test-day record of cow  $j$ , made on day  $t$ , in which  $k$  is the  $n^{\text{th}}$  parameter of coefficient of Legendre polynomials for the 4<sup>th</sup> or 5<sup>th</sup> order; and  $\varepsilon_{ijkl}$  is the random residual. Orthogonal Legendre polynomials were calculated as showed by Kirkpatrick et al. (1994).

It was assumed that:

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

in which:  $G$  and  $P$  are covariance matrices of the random regression coefficients;  $R = I\sigma_e^2$  is a diagonal matrix (residual);  $\otimes$  is a Kronecker product between matrices; and  $I$  is an identity matrix.

Records of 305-day fat and protein yields were used in single trait 305-day lactation models, named as LMF (fat) and LMP (protein), which included effects of herd-year-season of calving and age at calving (linear covariable) as fixed effects, and additive genetic animal and residual as random effects. The model used to estimate genetic parameters and EBVs for 305F and 305P was  $Y_{ij} = \text{HYS}_i + b_n \chi_{ij} + a_{ij} + \varepsilon_{ij}$ , in which:  $Y_{ij}$  is the 305-day fat or protein yield record of animal  $j$ , in herd-year-season of calving  $i$ ;  $\text{HYS}_i$  is the effect of herd-year-season of calving  $i$ ;  $b_n$  is the the linear regression coefficient for 305-day yield, as a function of age at calving (linear covariable);  $\chi_{ij}$  is the age of cow at calving, in months;  $a_{ij}$  is the additive genetic effect of animal  $j$  in herd-year-season of calving  $i$ ; and  $\varepsilon_{ij}$  is the residual effect.

The analyses were performed by REMLF90 software (Miszta 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:

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

in which:  $\hat{\alpha}_i$  is a vector ( $k_a \times 1$ ) of the estimates of additive genetic random regression coefficients specific to the animal  $i$ , and  $z'_t$  is a vector ( $k_a \times 1$ ) of Legendre polynomial coefficients evaluated at day  $t$ , which may be illustrated for a fifth-order Legendre polynomial:

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

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

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

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

The models were compared using the values of residual variance (RV), the Akaike's information criterion  $\text{AIC} = -2\log L + 2p$ , and the Schwarz's Bayesian information criterion  $\text{BIC} = 2\log L + p \log(\lambda)$ , where  $p$  is the number of parameters in the model. Using REML,  $\lambda = N - r(X)$ , in which:  $N$  is the number of test-day records;  $r(X)$  is the rank of the fixed effects incidence matrix; and  $2\log L$  is provided as default by REMLF90. The best model is indicated by the lowest values of AIC and BIC. A log-likelihood ratio test (LRT) was applied to test the significant differences between models with different orders of LP. The null hypothesis ( $H_0$ ) implied that restricted likelihood functions of the models did not differ when the number of parameters increased. The calculated value of LRT was compared to the chi-square Table ( $\chi^2$ ) with ten degrees of freedom, at 5% probability.

## Results and Discussion

AIC, BIC and 2LogL had the lowest values (highlighted) for RRF4 and RRP4 models (Table 1).

Residual values showed 4% (fat) and 5% (protein) decreasing, with an increasing in the order of Legendre polynomials. LRTs between RRF4 and RRF5, and between RRP4 and RRP5 were found to be different ( $p < 0.05$ ), and they varied from the fourth to the fifth-order of polynomials. A reduction of the values 2LogL, AIC, and BIC, as Legendre polynomial order decreases with a significant change in the log likelihood, indicated random regression models of the fourth-order Legendre polynomials as the best fit, but residual values indicated the fifth-order as the model. In some studies, AIC, BIC, and RV indicated models with larger number of parameters in the literature (Biassus et al., 2011; Aliloo et al., 2014). The present results may be a consequence of hyper parameterization of the models. AIC and BIC criteria favor simpler models because of the penalty term.

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 effects, 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%. According to 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. Although AIC, BIC, 2LogL, and residual variances showed conflicting results in the present study, the use of different approaches may help to explain the differences between random regression models. The analysis of the eigenvalues indicates the decreasing importance of adding more parameters (Aliloo et al., 2014). The first three additive eigenvalues explained a sufficiently large proportion of the variances for the models RRF4 and RRP4. The highest values of permanent environmental eigenvalues suggested that the permanent environmental effect could be modeled with a fifth-order Legendre polynomial, in comparison to the genetic effect. Although these results suggested different Legendre polynomial orders for additive genetic and permanent environmental effects, there is not a consensus in the literature (Araújo et al., 2006; Costa et al., 2008; Biassus et al., 2011; Aliloo et al., 2014). Aliloo et al. (2014) reporting that the

**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	3191455.3	3191497.3	3191569.1	11500		
RRF5	31	3345709.1	3345771.1	3345877.1	11040	154,273.8*	18.30
RRP4	21	3045851.8	3045893.8	3045965.5	5763		
RRP5	31	3199715.1	3199777.1	3199883.0	5469	153,917.5*	18.30

\*Significant at 5% probability. RRF4 and RRF5, random regression models for fat yield fitted by fourth and fifth-order Legendre polynomials; RRP4 and RRP5, random regression models for protein yield fitted by fourth and fifth-order Legendre polynomials; RV, residual value; LRT, likelihood ratio test.

**Table 2.** Eigenvalues ( $\lambda_i$ ) of the additive genetic variance and covariance matrix, and the respective percentual 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 models for fat yield fitted by fourth and fifth-order Legendre polynomials; RRP4 and RRP5, random regression models for protein yield fitted by fourth and fifth-order Legendre polynomials.

models with higher orders for genetic and permanent environmental effects were the best fit in Iranian Holsteins. However, Mohammadi & Alijani (2014), also using Holstein data in Iran, indicated lower orders of Legendre polynomials for genetic effects, and higher ones for permanent environmental effects. In Brazil, Araújo et al. (2006) and Biassus et al. (2011), who used the same orders for genetic and permanent environmental variances, 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.

For 305-day yield trait (Table 3), the heritabilities obtained from 305-day lactation models were 0.24 (LMF) and 0.17 (LMP); and random regression models showed heritability estimates of 0.21 (RRF4 and RRF5) and 0.20 (RRP4 and RRP5). The heritability estimates obtained from random regression models were very similar to those obtained from 305-day lactation models for 305-day fat and protein yields. In the literature, the estimates of heritability at 305 days, applying random regression models to analyze Holstein cattle, ranged from 0.29 to 0.41 for fat, and from 0.29 to 0.41 for protein yields (Bohmanova et al., 2008; Biassus et al., 2011; Kheirabadi & Alijani, 2014). Heritability estimates using 305-day lactation models were 0.13 for fat, and 0.12 for protein yield (Kim et al., 2009).

The heritability for test-day fat and protein yields, obtained from different random regression models, showed very similar trajectories on days in milk (Figure 1). The heritability of test-day values ranged from 0.13 to 0.23 for fat (RRF4 and RRF5), and

from 0.10 to 0.23 for protein (RRP4 and RRP5). Trajectories of additive genetic variances of test-day fat and protein yields were similar between different random regression models. Additive genetic variances of test-days ranged from 4,954.4 g<sup>2</sup> to 7,115.6 g<sup>2</sup> for fat (RRF4 and RRF5), and from 2,625.8 g<sup>2</sup> to 4,857.3 g<sup>2</sup> for protein (RRP4 and RRP5). Permanent environmental variances of test-days ranged from 12,136.6 g<sup>2</sup> to 32,909.8 g<sup>2</sup> for fat (RRF4 and RRF5 models), and from 9,093.8 g<sup>2</sup> to 19,702.0 g<sup>2</sup> for protein (RRP4 and RRP5). Heritability of the 305-day fat yield, obtained from 305-day lactation models, was similar to the values of heritability of test-day fat yield in mid-lactation, but it was lower than those at the extremes of lactation obtained from random regression models (RRF4 and RRF5). For protein yield, heritability of test-days, obtained from random regression models (RRP4 and RRP5), was higher in mid-lactation (90 to 270 days in milk) than that for 305-day yield estimated by 305-day lactation models. Kim et al. (2009) carried out a similar comparison and found quite similar results in a population of Holstein in Korea, and they reported that random regression models showed higher heritability of test-days on days in milk than the 305-day lactation models for fat and protein yields. In previous studies with a Brazilian Holstein population from Minas Gerais state, Biassus et al. (2011) found values ranging from 0.03 to 0.21 and 0.09 to 0.33 for fat and protein yields, respectively, estimated from random regression models. Rzewuska & Strabel (2013) and Abdullahpour et al. (2013) reported very close average heritability values from 0.17 to 0.22, and 0.14 to 0.23, for fat and protein yields, respectively. Trajectories of permanent environmental variances of test-day fat and protein yields were quite constant between 30 and 240 days in milk, in all models, and coincided with the highest heritability of test-day yields in mid-lactation. The lowest values of additive genetic and heritability of test-days, and the highest values of permanent environmental variances were found at the extremes of lactation curves. That pattern was also found by Biassus et al. (2011).

Spearman correlations of EBVs of bulls for 305-day yield, between 305-day lactation models and random regression models, increased from 0.86 to 0.97 (fat) and 0.86 to 0.98 (protein) with the increase in the classes of progeny size (Table 4). Spearman correlations of EBVs of cows for 305-day yield increased from 0.83 to 0.89

**Table 3.** Estimates of heritability of 305day fat and protein yields estimated from conventional and random regression models.

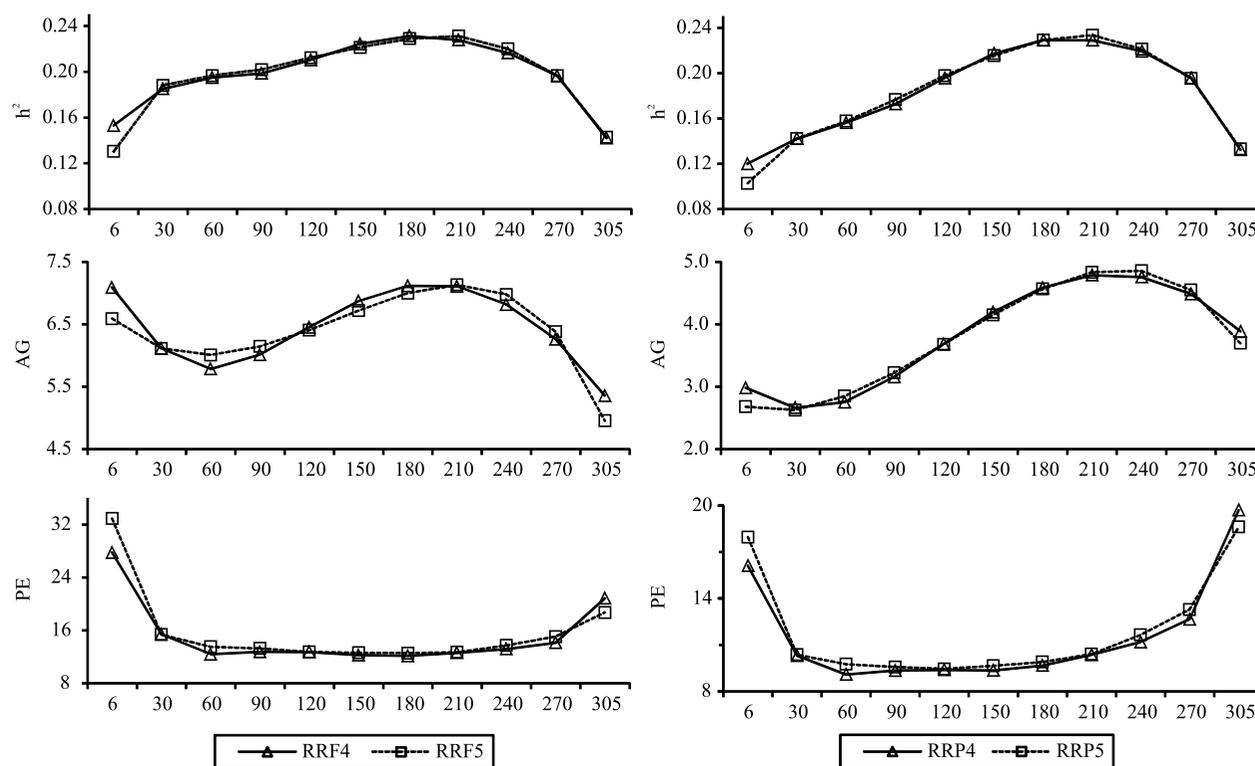
Trait	Models	h <sup>2</sup>
Fat	LMF	0.24
	RRF4	0.21
	RRF5	0.21
	LMP	0.17
Protein	RRP4	0.20
	RRP5	0.20

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

(fat) and 0.81 to 0.86 (protein) for groups of cows with 6 to 10 test-days. Those results of Spearman correlations of bulls may be interpreted in two different ways. On the one hand, Spearman correlations between EBVs estimated by LMF and RRF4, and between EBVs estimated by LMP and RRP4, increased with the increase in the information amount of the progeny size of bulls. It could suggest that random regression and 305-day lactation models might have shown very similar ranks of bulls, if a large amount of information was available. On the other hand, as bulls' progeny size and number of test-days of cows decreased, there was a decrease in the Spearman correlations of EBVs of bulls and cows. In these results, it became evident the increase in the re-ranking of animals, as the amount of information decreased.

The average gain in reliability of EBVs of bulls increased from 3 to 16%, when RRF4 were compared to LMF, and from 6 to 26%, when RRP4 were compared to LMP, according to the decrease in the classes of progeny size (Table 5). However, the increase in the number of test-days of cows did not show a pattern of

increase or decrease in the gain of reliability of EBVs, when random regression models were compared with 305-day lactation models. For cows, the average gain in reliability of EBVs ranged between 24 and 26%, when RRF4 was compared to LMF, and between 38 and 40%, when RRP4 was compared to LMP. Although the increase in the bulls' progeny size was followed by a decrease in the gain in reliability of EBVs of bulls, random regression models estimated EBVs with more reliability than the conventional 305-day lactation models. Considering bulls with progeny size class of 200 to 399, the gain in reliability of EBVs for 305-day yield ranged from 1 to 10% (in average 3%) for fat, and from 2 to 18% for protein (in average 6%). The highest gains in reliability of EBVs of bulls were found when classes of lower progeny size were considered. Considering bulls with classes of progeny size below 200 to 399, the average percentage of gain in reliability was from 5 to 16% for fat, and it was higher for protein – between 8 and 26%. Those gains in reliability ranged from 2 to 39% for fat, and from 3 to 63% for protein, when considering the range in parentheses, instead of



**Figure 1.** Heritability ( $h^2$ ), additive genetic (AG), and permanent environmental (PE) variances estimated from random regression models fitted by Legendre polynomials of fourth and fifth orders for fat (RRF4 and RRF5) and for protein (RRP4 and RRP5).

the average gain of the classes. The number of bulls decreased with the increase in the size of progeny size class, which means that some bulls have been used more intensively than others. The most relevant gain in reliability of EBVs was for bulls with lower number of progenies, which suggests that young bulls could have their EBVs estimated more precisely using a random

**Table 4.** Spearman 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 during lactation.

Progeny size	Number of bulls or cows	Models	
		RRF4	RRP4
Bulls			
200 to 399	29	0.97	0.98
100 to 199	74	0.96	0.94
50 to 99	154	0.95	0.92
25 to 49	175	0.90	0.87
10 to 24	352	0.86	0.86
Cows			
Number of test-days			
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 models for fat and protein yield fitted by fourth-order Legendre polynomial.

regression model. The average percentage of gain in reliability of EBVs for 305-day yield of cows between random regression and 305-day lactation models were similar for classes of 6, 7, 8, 9 or 10 test-days in lactation. It suggested that, in the conditions of the present study, a group of cows with lower number of test-days, as 6 for instance, may show a similar average reliability and a similar average gain in reliability as those of cows with complete lactation records, which are 10 test-days in lactation. The gain in reliability of breeding values in the genetic evaluations may bring benefits to selection-based breeding programs, if a larger number of cows and bulls is considered for genetic evaluations, instead of being deleted because of low number of records. In Brazil, one of the main reasons for the low number of recorded test-days of cows has been the increased recording costs, besides the low milk prices, and the lack of financial support from the government.

EBVs for 305-day fat and protein yields by year of birth of cows and bulls showed positive trends ranging between 0.16 and 0.42 kg per year (RRF4 and LMF), and between 0.47 and 0.76 kg per year (RRP4 and LMP), from cows born from 1990 to 2008, and for bulls born from 1979 to 2001 (Table 6). All trends of EBVs by year of birth of bulls were significantly different from zero at  $p < 0.01$ , except for RRF4 model; for cows, all trends

**Table 5.** Average percentage of gain in reliability (range in parenthesis) of EBVs for 305-day yield estimated by random regression models, in comparison to 305-day lactation models.

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

<sup>(1)</sup>Average and standard-deviation of EBV reliability values of 305-day yield estimated from 305-day lactation models. <sup>(2)</sup>Average percentage of proportional gains of reliability compared to 305-day lactation models 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.

**Table 6.** Estimates of breeding value trends (b) of bulls and cows by year of birth, standard errors (SE), and coefficients of determination ( $R^2$ ), for cows born between 1990 and 2008, and bulls from 1979 to 2001.

Trait	Bulls			Cows		
	b	SE	$R^2$	b	SE	$R^2$
LMF	0.31*	0.09	0.33	0.42**	0.06	0.78
LMP	0.47**	0.06	0.76	0.52**	0.03	0.97
RRF4	0.16 <sup>ns</sup>	0.10	0.10	0.34**	0.06	0.63
RRP4	0.65**	0.07	0.77	0.76**	0.04	0.95

LMF, 305-day lactation models 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. \*,\*\*Significant at 1 and 0.1%, respectively.

were significantly different from zero at  $p < 0.001$ . Genetic trend values in the literature ranged between 0.25 and 0.60 kg per year for 305-day fat yield, and around 0.45 for 305-day protein yields (Durães et al., 2001; Sawalha et al., 2005). The genetic trends in the present study accounted for approximately 0.07% (fat) and 0.18% (protein) of the average phenotypic yield, which implies that fat and protein yields have improved slightly in Brazilian Holsteins. Such genetic trends might be attributed to the emphasis on the selection for increased milk yield by the Brazilian dairy industry. Moreover, the Brazilian Holstein population structure is based on imported genetic material from the USA, Europe, and Canada (Silva et al., 2016). It is a fact that breeding programs in Brazil, Canada or Europe have different objectives of selection.

## Conclusions

1. The adoption of a random regression model using Legendre polynomials will increase the reliability of estimated breeding values in the genetic evaluations.

2. A random regression model using the fourth-order Legendre polynomial is the most recommended model for genetic evaluations of fat and protein yields of Brazilian Holstein cattle.

## Acknowledgements

To Embrapa Gado de Leite and to Associação Brasileira de Criadores de Bovinos da Raça Holandesa (ABCBRH), for providing the data for analyses; to Dr. Claudio Napolis Costa, for his valuable contributions;

and to Centro Nacional de Processamento de Alto Desempenho em São Paulo (Cenapad, SP), for supporting this research in its data processing phase.

## References

- ABDULLAHPOUR, R.; MORADI SHAHRBABA, M.; NEJATIHAVAREMI, A.; VAEZTORSHIZI, R.; MRODE, R. Genetic analysis of milk yield, fat and protein content in Holstein dairy cows in Iran: Legendre polynomials random regression model applied. *Archiv Tierzucht*, v.56, p.497-508, 2013. DOI: 10.7482/0003943856048.
- ALILOO, H.; MIRAIEASHTIANI, S.R.; SHAHRBABA, M.M.; URIOSTE, J.I.; SADEGHI, M. Accounting for heterogeneity of phenotypic variance in Iranian Holstein test-day milk yield records. *Livestock Science*, v.167, p.25-32, 2014. DOI: 10.1016/j.livsci.2014.05.012.
- ARAÚJO, C.V. de; TORRES, R. de A.; COSTA, C.N.; TORRES FILHO, R. de A.; ARAÚJO, S.I.; LOPES, P.S.; REGAZZI, A.J.; PEREIRA, C.S.; SARMENTO, J.L.R. Uso de modelos de regressão aleatória para descrever a variação genética da produção de leite na raça Holandesa. *Revista Brasileira de Zootecnia*, v.35, p.975-981, 2006. Suplemento. DOI: 10.1590/S151635982006000400006.
- BANGA, C.B.; NESER, F.W.C.; GARRICK, D.J. Breeding objectives for Holstein cattle in South Africa. *South African Journal of Animal Science*, v.44, p.199-213, 2014. DOI: 10.4314/sajas.v44i3.1.
- BIASSUS, I. de O.; COBUCI, J.A.; COSTA, C.N.; RORATO, P.R.N.; BRACCINI NETO, J.; CARDOSO, L.L. Genetic parameters for production traits in primiparous Holstein cows estimated by random regression models. *Revista Brasileira de Zootecnia*, v.40, p.85-94, 2011. DOI: 10.1590/S151635982011000100012.
- BIASSUS, I. de O.; COBUCI, J.A.; COSTA, C.N.; RORATO, P.R.N.; BRACCINI NETO, J.; CARDOSO, L.L. Persistence in milk, fat and protein production of primiparous Holstein cows by random regression models. *Revista Brasileira de Zootecnia*, v.39, p.2617-2624, 2010. DOI: 10.1590/S151635982010001200009.
- BIGNARDI, A.B.; EL FARO, L.; CARDOSO, V.L.; MACHADO, P.F.; ALBUQUERQUE, L.G. Random regression models to estimate test-day milk yield genetic parameters Holstein cows in Southeastern Brazil. *Livestock Science*, v.123, p.17, 2009. DOI: 10.1016/j.livsci.2008.09.021.
- BIGNARDI, A.B.; EL FARO, L.; TORRES JÚNIOR, R.A.A.; CARDOSO, V.L.; MACHADO, P.F.; ALBUQUERQUE, L.G. Random regression models using different functions to model test-day milk yield of Brazilian Holstein cows. *Genetics and Molecular Research*, v.10, p.3565-3575, 2011. DOI: 10.4238/2011.October.31.4.
- BOHMANOVA, J.; MIGLIOR, F.; JAMROZIK, J.; MISZTAL, I.; SULLIVAN, P.G. Comparison of random regression models with Legendre polynomials and linear splines for production traits and somatic cell score of Canadian Holstein cows. *Journal of Dairy Science*, v.91, p.3627-3638, 2008. DOI: 10.3168/jds.20070945.

- COBUCI, J.A.; COSTA, C.N.; NETO, J.B.; FREITAS, A.F. de. Genetic parameters for milk production by using random regression models with different alternatives of fixed regression modeling. **Revista Brasileira de Zootecnia**, v.40, p.557-567, 2011. DOI: 10.1590/S151635982011000300013.
- COSTA, C.N.; CARVALHEIRA, J.; COBUCI, J.A.; FREITAS, A.F.; THOMPSON, G. 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**, v.39, p.165-168, 2009. Supplement 1.
- COSTA, C.N.; MELO, C.M.R. de; PACKER, I.U.; FREITAS, A.F. de; TEIXEIRA, N.M.; COBUCI, J.A. Genetic parameters for test day milk yield of first lactation Holstein cows estimated by random regression using Legendre polynomials. **Revista Brasileira de Zootecnia**, v.37, p.602-608, 2008. DOI: 10.1590/S151635982008000400003.
- DEGROOT, B.J.; KEOWN, J.F.; VAN VLECK, L.D.; KACHMAN, S.D. Estimates of genetic parameters for Holstein cows for test-day yield traits with a random regression cubic spline model. **Genetics and Molecular Research**, v.6, p.434-444, 2007.
- DURÃES, M.C.; FREITAS, A.F.de; VALENTE, J.; TEIXEIRA, N.M.; BARRA, R.B. Tendência genética para a produção de leite e de gordura em rebanhos da raça holandesa no Estado de Minas Gerais. **Revista Brasileira de Zootecnia**, v.30, p.66-70, 2001. DOI: 10.1590/S151635982001000100011.
- FERREIRA, W.J.; TEIXEIRA, N.M.; EUCLYDES, R.F.; VERNEQUE, R. da S.; LOPES, P.S.; TORRES, R. de A.; WENCESLAU, A.A.; SILVA, M.V.G.B. da; MAGALHÃES JÚNIOR, M.N. Avaliação genética de bovinos da raça Holandesa usando a produção de leite no dia do controle. **Revista Brasileira de Zootecnia**, v.32, p.295-303, 2003. DOI: 10.1590/S151635982003000200007.
- JENSEN, J. Genetic evaluation of dairy cattle using test-day models. **Journal of Dairy Science**, v.84, p.2803-2812, 2001. DOI: 10.3168/jds.S00220302(01)747364.
- KHEIRABADI, K.; ALIJANI, S. Genetic parameters for milk production and persistency in the Iranian Holstein population by the multitrait random regression model. **Archiv Tierzucht**, v.57, p.112, 2014.
- KIM, B.W.; LEE, D.; JEON, J.T.; LEE, J.G. 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 Sciences**, v.22, p.923-930, 2009. DOI: 10.5713/ajas.2009.80110.
- KIRKPATRICK, M.; HILL, W.G.; THOMPSON, R. Estimating the covariance structure of traits during growth and ageing, illustrated with lactation in dairy cattle. **Genetics Research**, v.64, p.57-69, 1994. DOI: 10.1017/S0016672300032559.
- MELO, C.M.R. de; PACKER, I.U.; COSTA, C.N.; MACHADO, P.F. Genetic parameters for test day milk yields of first lactation Holstein cows by random regression models. **Animal**, v.1, p.325-334, 2007. DOI: 10.1017/S1751731107685036.
- MISZTAL, I.; TSURUTA, S.; LOURENÇO, D.; AGUILAR, I.; LEGARRA, A.; VITEZICA, Z. **Manual for BLUPF90 family of programs**. Available at: <[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 on: Oct. 30 2014.
- MISZTAL, I.; WIGGANS, G.R. Approximation of prediction error variance in largescale animal models. **Journal of Dairy Science**, v.71, p.27-32, 1988. DOI: 10.1016/S00220302(88)799762.
- MOHAMMADI, A.; ALIJANI, S. Estimation of genetic parameters and comparison of random regression animal and sire models of production traits in the first three lactations of Iranian Holsteins. **Biotechnology in Animal Husbandry**, v.30, p.261-279, 2014. DOI: 10.2298/BAH1402261M.
- MUIR, B.L.; KISTEMAKER, G.; JAMROZIK, J.; CANAVESI, F. Genetic parameters for a multiple-trait multiple-lactation random regression test-day model in Italian Holsteins. **Journal of Dairy Science**, v.90, p.1564-1574, 2007. DOI: 10.3168/jds.S00220302(07)716429.
- RZEWUSKA, K.; STRABEL, T. Genetic parameters for milk urea concentration and milk traits in Polish Holstein-Friesian cows. **Journal of Applied Genetics**, v.54, p.473-482, 2013. DOI: 10.1007/s1335301301598.
- SAWALHA, R.M.; KEOWN, J.F.; KACHMAN, S.D.; VAN VLECK, L.D. Genetic evaluation of dairy cattle with test-day models with autoregressive covariance structures and with a 305d model. **Journal of Dairy Science**, v.88, p.3346-3353, 2005. DOI: 10.3168/jds.S00220302(05)730186.
- SCHAEFFER, L.R.; JAMROZIK, J.; KISTEMAKER, G.J.; VAN DOORMAAL, B.J. Experience with a test-day model. **Journal of Dairy Science**, v.83, p.1135-1144, 2000. DOI: 10.3168/jds.S00220302(00)749794.
- SILVA, M.H.M.A. da; MALHADO, C.H.M.; COSTA Jr., J.L.C.; COBUCI, J.A.; COSTA, C.N.; CARNEIRO, P.L.S. Population genetic structure in the Holstein breed in Brazil. **Tropical Animal Health and Production**, v.48, p.331-336, 2016. DOI: 10.1007/s1125001509567.
- YAMAZAKI, T.; HAGIYA, K.; TAKEDA, H.; SASAKI, O.; YAMAGUCHI, S.; SOGABE, M.; SAITO, Y.; NAKAGAWA, S.; TOGASHI, K.; SUZUKI, K.; NAGAMINE, Y. 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**, v.152, p.120-126, 2013. DOI: 10.1016/j.livsci.2012.12.015.

---

Received on March 23, 2016 and accepted on July 21, 2016