Single-step genomic evaluation using an autoregressive multiple lactation model for Portuguese Holstein cattle

A.A. Silva[1], D.A. Silva[1], S.E.F. Guimarães[1], P.S. Lopes[1], F.F. Silva[1], G. Thompson[2,3], C.N. Costa[4] & J. Carvalheira[2,3]

[1] *Universidade Federal de Viçosa, Departamento de Zootecnia, 36570-900 Viçosa, MG, Brazil*
silvaalessandrazb@gmail.com *(Corresponding Author)*
[2] *Research Center in Biodiversity and Genetic Resources (CIBIO-InBio), Univ. of Porto, 4485-661 Vairão, Portugal*
[3] *Institute of Biomedical Sciences Abel Salazar (ICBAS), Univ. of Porto, Porto, Portugal*
[4] *Embrapa Gado de Leite, 36.038-330 Juiz de Fora, MG, Brazil*

Summary

This study aims to evaluate the feasibility of using the single-step GBLUP (ssGBLUP) to analyze milk yield using an autoregressive (AR) test-day model for Holstein cattle in Portugal. A total of 11,168,530 test-day (TD) records from the first three lactations of milk, from 1994 to 2016 were provided by the Portuguese Dairy Cattle Breeders Association. Genotypes of 1,081 bulls were also available. The data set were edited according to predefined criteria for genetic analysis with the AR model. Validation of the ssGBLUP was based on comparisons between models using the traditional pedigree-based additive genetic relationship matrix (A-AR) and the H relationship matrix, which included both, genomic and pedigree information (H-AR). Correlations between DYD from the complete data set and GEBV or parent average (PA) from a reduced data, in which daughters (12,586) of genotyped bulls (110) born after 2004 were excluded, were used to evaluate the predictive ability of the ssGBLUP. There were no practical differences between EBV rankings (r > 0.99) and genetic trends obtained from A-AR and H-AR. The relatively small number of genotyped animals may explain this result. The largest impact on reliabilities (R^2) was obtained on bulls with less than 10 daughters with average (minimum) of 0.59 (0.24) using A-AR vs. 0.66 (0.32) using H-AR. The correlation between GEBV and DYD of the 110 bulls with no daughter information on the reduced data set was 0.38 for H-AR. In the traditional evaluation, the correlation between PA and DYD was 0.26 for A-AR. When moving from A-AR to H-AR, the correlation increases 12%. These results suggest that the ssGBLUP autoregressive test-day model is feasible and may be applied on National Portuguese genetic evaluations. With the increase of genotyped animals in the future, including females, it is expected that the H-AR will provide higher R^2 especially on young bulls and improve the genetic progress of the Portuguese Holstein population.

Keywords: autoregressive test-day model, Holstein cattle, single-step GBLUP

Introduction

Genomic selection using single-nucleotide polymorphism (SNP) has become a standard procedure in dairy cattle breeding because of it's potential for increasing genetic gain through improved evaluation accuracy and by reducing generation interval. The current method used in dairy cattle evaluations is based on a multistep procedure (VanRaden *et al.* 2008). In the multi-step approach, the calculation of pseudo-phenotypes such as de-regressed EBV or daughter yield deviations (DYD) is required (Aguilar *et al.* 2010; Misztal *et al.* 2009).

N~ (0,), where I_L is the identity matrix within lactation and is the residual variance within lactation, and and are incidence matrices relating observations to fixed and random effects. A first order autoregressive covariance structure was assumed for HTD (within herds), LTE (between parities) and STE (between TD, within lactations) effects. In the **H** matrix, the original genomic relationship matrix (**G**) was constructed according to VanRaden (2008). Variance components were estimated using DFREML procedures following assumptions of the AR model. To evaluate the predictive ability of the ssGBLUP, correlations were obtained between DYD from the complete data set and GEBV or parent average (PA) from a reduced data, in which daughters (12,586) of genotyped bulls (110) born after 2004 were excluded.

Results and Discussion

Estimated genetic and residual variances were smaller for the H-AR model compared to the A-AR model, resulting in similar heritabilities. There were no practical differences between the animal EBV rankings (r > 0.99) and the genetic trends were almost coincidental for the two models.

The impact of the inclusion of the genomic relationship is well depicted in Figure 1. As expected, the bulls that benefited more in terms of reliability (R^2) were the younger ones, with fewer daughters (< 10) reflecting a reduction on the prediction error variances. The average (minimum) R^2 obtained for this group of bulls, were 0.66 (0.32) with H-AR compared with 0.59 (0.24) for A-AR, respectively (Figure 1-A). On the other hand, no expressive changes were observed in R^2 in non-genotyped bulls or genotyped bulls with more than 10 daughters (Figure 1-B). This is not consistent with previous studies by Guo *et al.* (2015) on pigs, which reported improved reliabilities for non-genotyped animals. Most probably, this is due to the relatively small number of genotyped bulls in the present study.
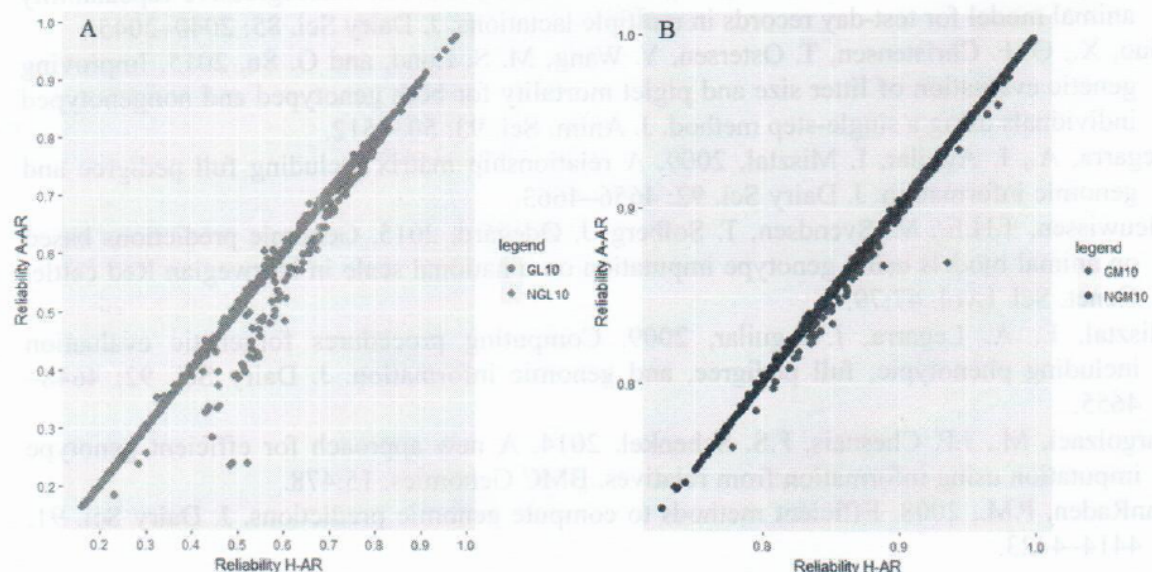

Figure 1. Reliabilities of EBV (R^2) obtained using H-AR and A-AR models. A- Reliabilities for genotyped (GL10) and non-genotyped (NGL10) bulls with less than 10 daughters. B- Reliabilities for genotyped (GM10) and non-genotyped (NGM10) bulls with more than 10 daughters.

Random vs fixed contemporary group effects on genetic evaluations of the Brazilian Holstein cattle using an autoregressive animal model

D.A. Silva[1], A.A. Silva[1], C.N. Costa[2], F.F. Silva[1], P.S. Lopes[1], S.E.F. Guimarães[1], G. G. Santos[2], G. Thompson[3,4] & J. Carvalheira,[3,4,*]

[1] *Universidade Federal de Viçosa, Departamento de Zootecnia, 36570-000 Viçosa-MG, Brazil*
[2] *Embrapa Gado de Leite, 36.038-330 Juiz de Fora, MG, Brazil*
[3] *Research Center in Biodiversity and Genetic Resources (CIBIO-InBio), Univ. of Porto, 4485-661 Vairão, Portugal*
[4] *Institute of Biomedical Sciences Abel Salazar (ICBAS), Univ. of Porto, Porto, Portugal*
* *jgc3@cibio.up.pt (Corresponding Author)*

Summary

Test-day milk yields and somatic cell score (SCS) of Brazilian Holstein cattle were used to compare the effect of random vs fixed contemporary groups (herd-test-date, HTD) using an autoregressive testy-day animal model for genetic evaluations. Fitting the HTD effect as fixed may minimize prediction bias, but requires a minimum number of observations per contemporary group (CG) to simultaneously maximize the effective number of observations and minimize the residual error and prediction error variance. Because most dairy herds in Brazil are characterized by their small size (many CG have less than 3 observations), an alternative is to consider HTD as a random effect especially if non-random associations with sires are considered negligible. A total of 3,428,352 and 2,853,695 testy-day records from the first three lactations of milk yield and somatic cell scores (SCS), respectively, were used for this study. This data was provided by the Brazilian Holstein Cattle Breeders Association and obtained from 1994 to 2016. The additive genetic variances estimated with random CG increased in both traits but also did the estimated phenotypic variances, leaving the heritabilities almost the same. Nevertheless, the annual genetic gains for milk yield (SCS), favored the random CG approach where for males, the yearly progress was 42 kg (-0.034 scores) greater (smaller) than when the fixed CG was fitted. Similarly, for females, the yearly progress was 27 kg (-0.002 scores) greater (smaller) than when the fixed CG was fitted to the model. Steeper genetic trends with random CG indicate that younger animals tended to have higher EBVs giving emphasis to the selection program. These results indicate that using CG as a random effect within the autoregressive animal model should be appropriate for the genetic evaluations of milk yield and SCS for the Brazilian Holstein cattle.

Keywords: autoregressive animal model, contemporary groups, genetic trends

Introduction

In an animal model genetic evaluation framework, the minimum number of observations within each contemporary group (CG) is a critical factor to reduce bias on the prediction of breeding values. Fitting CG as fixed or random is a matter of balance between accuracy and bias. Using the random approach may recover some information across CG (Visscher and Goddard, 1993; Strabel et al., 2005) especially when the number of observations in some CG is small, but it will introduce bias to the breeding value predictions if there are non-random associations of animals with effects considered to be fixed (Vasconcelos *et al.*, 2008). If these

are shown in Table 1. The additive genetic variances estimated with random CG increased in the two traits but also did the estimated phenotypic variances, leaving the heritabilities almost the same (Table 1). Costa *et al.* (2014), also using an AR model with fixed CG, analyzed Brazilian Holstein data but restricted the size of the CG to those with 3 or more observations. They found slightly better estimates for heritabilities (0.23 to 0.35 and 0.19 to 0.20 for milk yield and SCS, respectively). The estimated variance associated with the animal non-genetic component was especially important between TD within lactations (STE), revealing a strong permanent environmental effect at this stage. This pattern may be expected from breeds with high genetic potential, producing in high input production systems (Carvalheira *et al.*, 2002).

Table 1. Estimates of variance components, autocorrelations and genetic parameters for milk yield and somatic cell score (SCS) for Brazilian Holstein cattle, assuming contemporary groups as random (HTDR) or fixed (HTDF) effects in AR models.

Estimates[1]	HTDF		HTDR	
	Milk ± SE	SCS ± SE	Milk ± SE	SCS ± SE
	10.32 ± 0.113	0.57 ± 0.010	11.03 ± 0.115	0.60 ± 0.012
	<0.001 ± <0.0001	<0.001 ± <0.0001	<0.001 ± <0.0001	<0.001 ± <0.0001
	<0.001 ± <0.0001	<0.001 ± <0.0001	<0.001 ± <0.0001	<0.001 ± <0.0001
	11.74 ± 0.100	1.08 ± 0.019	11.68 ± 0.095	1.08 ± 0.019
	14.49 ± 0.215	1.18 ± 0.014	14.39 ± 0.214	1.18 ± 0.013
	14.96 ± 0.225	1.12 ± 0.012	14.87 ± 0.227	1.12 ± 0.012
	21.31 ± 0.321	1.46 ± 0.021	21.26 ± 0.325	1.46 ± 0.022
	0.69 ± 0.002	0.85 ± 0.002	0.69 ± 0.002	0.85 ± 0.002
	29.38 ± 0.578	1.83 ± 0.036	29.43 ± 0.557	1.83 ± 0.037
	0.78 ± 0.002	0.83 ± 0.001	0.78 ± 0.002	0.82 ± 0.001
	36.26 ± 0.729	2.32 ± 0.044	36.66 ± 0.721	2.31 ± 0.043
	0.76 ± 0.001	0.84 ± 0.002	0.76 ± 0.001	0.84 ± 0.002
	-	-	3.19 ± 0.139	0.27 ± 0.012
	-	-	0.72 ± 0.012	0.63 ± 0.012
	0.24 ± 0.004	0.18 ± 0.002	0.23 ± 0.002	0.18 ± 0.002
	0.19 ± 0.004	0.16 ± 0.002	0.19 ± 0.003	0.15 ± 0.002
	0.17 ± 0.003	0.14 ± 0.002	0.17 ± 0.002	0.14 ± 0.002

[1] : additive genetic variance (kg^2 for milk and score units^2 for SCS); : LTE variance; : STE variance; : herd-test-day (HTD) variance; : residual variance; , e : autocorrelations for LTE, STE and HTD; : heritability, where =1, 2 and 3 correspond to first, second and third lactations, respectively.

The influence of random and fixed CG effects on the genetic trends of milk yield and SCS are illustrated in Figure 1. Assuming the CG as random had an enormous impact on the estimated genetic progress of the Brazilian Holstein cattle. A much steeper trend was estimated for the HTDR model in both traits. Annual genetic gains for milk yield were 82.14 kg/yr. compared with 39.95 kg/yr. for cows and 57.96 kg/yr. compared with 30.94 kg/yr. for bulls, respectively. For SCS, annual genetic gains (score units/yr.) were -0.037 *vs.* -0.003 and -0.022 *vs.* -0.020 for cows and bulls, respectively.

Adjustment of Lactation Curves of Holstein Cows from Herds of Minas Gerais, Brazil

Jairo Azevedo Junior[1], Tarcisio de Moraes Gonçalves[1], José Camisão de Souza[1], Mary Ana Petersen Rodriguez[2], Cláudio Napolis Costa[3] & Julio Gil Vale Carvalheira[4]

[1] Department of Animal Science, Federal University of Lavras, Lavras, MG, Brazil

[2] Department of Agrarian Sciences, State University of Montes Claros, Montes Claros, MG, Brazil

[3] Embrapa Gado de Leite, Juiz de Fora, MG, Brazil

[4] CIBIO, University of Porto, Porto, Portugal

Correspondence: Jairo Azevedo Junior, Department of Animal Science, Federal University of Lavras, Lavras, MG, Brazil. Tel: 55-63-999-336-010. E-mail: jairo@zootecnista.com.br

Abstract

Random regression models (RRM) differ in terms of the functions used to describe the shape of lactation curves. The aim was to compare random regression models under different functions to describe the lactation curves from Holstein cows in herds of the state of Minas Gerais. A database of 28,118 production records was analyzed using the test-day records of 4,230 first parity cows from five herds. The Wilmink, Ali & Schaeffer and Legendre polynomial (orders 4, 5 and 6) functions were adjusted in RRM to model the mean production trend (fixed) and genetic and permanent environmental (random) effects. The residual variances were assumed to be constant throughout lactation. Analyses were performed using the AIREMLF90 program. Except for the model with the polynomial function of order 5, all models converged. The Wilmink function showed lower values for criteria based on the -2log (L), AIC and BIC. The model with the Legendre polynomial of order 6 showed lower residual variance. Heritability estimates were similar between functions, ranging from 0.07 to 0.18 and were higher from 215 days of lactation. From 155 days of lactation, genetic and permanent environmental correlations between successive controls are of high magnitude. The Wilmink function is the most suitable for the study of milk yields from primiparous Holstein cows. The selection of animals is possible from 155 days of lactation on. Permanent environmental effects have greater influence on the milk production at the end of lactation of primiparous cows and should be considered since they are important and may be cumulative throughout lactation.

Keywords: dairy cattle, heritability, models comparison, permanent environmental effects, random regression model

1. Introduction

The Holstein breed has the largest number of individuals in the world and has been, and still is, the target of an intensive genetic selection, which has turned it into an extremely productive animal.

According to the Foreign Agricultural Service of the US Department of Agriculture (USDA) December 2012 report on milk yield, Brazil was ranked third in the world. Minas Gerais is the Brazilian State with the highest annual milk production, with over 27% of the total production (Poll, 2013).

Each country and breed association defines a method for evaluating phenotypic and genetic traits of economic interest that usually are selected simultaneously, but with different weights, within selection indices.

Among the different milk yield analysis approaches throughout lactation, Random Regression Models (RRM) have numerous advantages in the genetic evaluation of dairy cattle.

In RRM, by the use of functions that define the production on day "t" as a linear function of time or of a logarithmic or exponential function of time, an adjustment of one global mean curve for the population is sought. Specific curves for each animal coincide, deviating in either direction from the overall curve, as the parameters of the individual curves are positioned, distributively, around the curve of the global parameters. Thus, through RRM, the genetic merit of each individual can be obtained at any point of lactation. In addition, the accuracies

method, based on normal distribution assumptions was adopted. Thus, the mean (μ) and standard deviation of the studied traits were obtained. Data within the interval μ±3 standard deviations accounted for 99.73% of the observations. Records outside this range were considered outliers (Lethen, 1996).

Firstly, the elimination of outliers was performed AC and subsequently for TDY. After elimination of the outliers, only lactations with at least three milk records were considered. A total of 28,118 milk production records from 4,230 animals were considered.

2.2 Genetic Analyses

The random regression model used to obtain the fixed and random solutions was:

$$y = HYM + HAGE + \sum bz + \sum az + \sum pz + e \tag{1}$$

where, y = test-day milk yield; HYM the fixed effect of the 401 classes comprising the effects of herd and date (year and month) of test day; $HAGE$-fixed effect of the 39 classes comprising the effects of herd and age at calving; b-fixed regression coefficients obtained for each herd; a-genetic random regression coefficients for each animal; p-random regression coefficients for permanent environmental effect for each animal; and e-random residual effect of each observation. The residual variance was assumed to be constant throughout the lactation and independent between milk records (Costa et al., 2008; Biassus, Cobuci, & Costa, 2011). The different functions adopted, characterizing the different sub-models of random regression, have been adjusted in $\sum bz$, $\sum az$ and $\sum pz$ in order to model the fixed and random animal effects and animal environmental permanent, respectively.

The functions of Wilmink (WIL), Ali & Schaeffer (ALI) and Legendre Orthogonal Polynomials of orders 4 to 6 (indicated by the polynomial degree, *e.g.* LEG3, LEG4 and LEG5, respectively) were adjusted by the program AIREMLF90 (Misztal et al., 2002).

The Wilmink function (1987) is represented by:

$$Yt = a + bt + ce^{-kt} \tag{2}$$

in which the parameters: Yt is the test day milk yield at t days of lactation; a is associated with milk production in early lactation; b to the increase in production until the lactation peak, the higher the magnitude of this parameter the greater the production before the peak; c is related to the decline in production after the peak, where higher values may be related to greater persistence; e means exponential (e is not a parameter to be estimated); and k is related to the peak moment. The k parameter is related to the peak moment and usually takes a fixed value that can be obtained from a preliminary analysis, which implis that the model has only three estimable parameters. However, Schaeffer et al. (2000) found no significant differences between the actual and the estimated productions, when the value of k is estimated for each situation analyzed. Thus, it was assumed that $k = 0.05$, associated with a peak occurring near 50 days of lactation, according to Wilmink (1987) and Schaeffer et al. (2000).

The Ali and Schaeffer (1987) function is represented by:

$$Yt = a + b\gamma_t + c\gamma_t^2 + d\omega_t + e\omega_t^2 \tag{3}$$

Where: $\gamma_t = (t/305)$, $\omega_t = \ln(305/t)$, Yt is the test day milk yield at t days of lactation; a is the production at the lactation peak; d and e are related to the rate of increase in milk production before the peak, while b and c are related to the decline in production after the lactation peak. When d and e have estimates of high magnitude, milk yield can reach its maximum (peak) in a few weeks after calving. The rate of decrease in production described by the parameters b and c can provide information about ability of the animal to maintain a fairly constant milk production (persistency of lactation) after the peak.

The Legendre Orthogonal Polynomials are a methodology published by Spiegel (1971) and modified by Kirkpatrick, Lofsvold, and Bulmer (1990). They are polynomial functions of n degree and domain $n + 1$, implying in the existence of an additional parameter other than the polynomial degree being estimated. In these functions, a single observation can be written as:

$$y_t = \sum_{i=0}^{n} \alpha_i \varnothing_1(\omega) \tag{4}$$

where, ω is a standardized time unit ranging from -1 to +1 (Kirkpatrick, Lofsvold, & Bulmer, 1990) and can be obtained by:

$$\omega = 2[(t - t_{min})/(t_{max} - t_{min})] - 1 \tag{5}$$

In this case, it is assumed that $t_{min} = 5$ days and $t_{max} = 305$ days, corresponding to the first and last test day after calving, respectively.

Table 1. Population description of five Holstein herds from in Minas Gerais, Brazil

Information	Frequency
Number of animals	8 652
Number of cows	8 202
Number of bulls	450
Number of unrelated animals	5 189

The high number (59.97%) of unrelated individuals (Table 1) was due to the failure in identifying in the records the complete genealogy. Sire and dam could not be identified in 4,645 (53.69%) records; Sire only was identified in 194 (2.24%) records; dam only was identified in 247 (2.85%) and in 3,566 (41.22%) sire and dam were known.

Only three individuals were inbred. The highest coefficient of inbreeding (F) observed was 25% and the average inbreeding coefficient of the three inbred animals was 20.83%. Possibly, the low number of detected inbred animals was the result of incomplete genealogy records.

A higher number of lactations were observed in herd 2 (Table 2). Nevertheless, in all herds, milk yield (TDY) was recorded in the total course of the lactation period. This fact allows a better fit of the functions to the test day data, since information from all the different stages of lactation was present (upswing, peak and decreasing production phase).

Table 2. Numbers of monthly milk production records (TD) and numbers of lactations from five Holstein herds in the state of Minas Gerais, Brazil, assessed in the first lactation

	Herd 1	Herd 2	Herd 3	Herd 4	Herd 5
TD 1	84	1,684	353	254	83
TD 2	91	1,883	383	270	94
TD 3	92	1,953	397	289	100
TD 4	103	1,968	383	283	97
TD 5	88	2,017	385	276	92
TD 6	105	2,095	370	289	108
TD 7	100	2,077	361	278	94
TD 8	99	2,159	353	283	90
TD 9	96	2,136	340	266	88
TD 10	77	1,966	283	228	75
Total TD	935	19,938	3,608	2,716	921
Lactations	147	2,999	508	436	140

For each animal, on average, 6.64 milk records were evaluated by lactation. Whereas the model with the most parameterized function (LEG5), the number of dairy controls registered for lactation may be considered sufficient for the adjustment of the evaluated functions.

The means and standard deviations of days in milk at the time of milk recording among the herds are similar (Table 3). Also, within herds, there is regularity in the milk record intervals. Such considerations are important because they show fairly similar conditions between the different herds sampled. Thus, more reliable comparisons of milk production curve patterns between herds are possible. In addition, the variations due to genetic and environmental effects that occur at specific times in the course of lactation may be quantified equally among the herds studied.

Apparent differences in productivity between herds and milk record data are observed on Table 4 and Figure 1. These are milk records from 4,230 lactations. Cows were 26.99 month-old on average (18.85 up to 47.78 months) at first calving.

Herd 1 had numerically lower milk production records compared to the remaining herds, and registered the highest production at test day in third month (TD3) (Table 4 and Figure 2), which was close to 70 days of lactation (Table 3).

Convergence was reached for all models studied, except for the adjusted model with LEG4 function, even after extending the number of iterations up to 200,000 cycles for this function.

Residual variances estimated in this study ranged from 14.11 for the model with function adjustment LEG5 to 17.56 for the model adjusted with the WIL function (Table 5). The lower values obtained for the residual variance models are associated with a greater number of parameters to be estimated, as it is observed when the model is adjusted with the LEG5 function (Table 5).

Table 5. Choice criteria for models fitted with different functions to data of test day milk yield from primiparous Holstein cows from five herds in the state of Minas Gerais, Brazil

Function	p	-2 log L	AIC	BIC	(σ_e^2)
WIL	13	180,259.80	180,285.80	180,317.64	17.56
LEG3	21	182,383.32	182,425.32	182,476.75	15.88
ALI	31	183,970.81	184,032.81	184,108.73	14.62
LEG4 (NC)	31	184,292.42	184,354.42	184,430.34	14.85
LEG5	43	186,636.05	186,722.05	186,827.36	14.11

Note. AIC = Akaike Information Criteria; ALI = Ali & Schaeffer; BIC = Schwarz Bayesian Criteria; LEG3 = Legendre polynomials degree 3, order 4; LEG4 = Legendre polynomials degree 4, order 5; LEG5 = Legendre polynomials degree 5, order 6; NC = Not convergence; WIL = Wilmink; p = Number of parameters of model; (σ_e^2) = residual variance.

Higher values for the information criteria [-2log (L), AIC and BIC] were observed for the most parameterized models, especially the BIC criterion, which is directly related to the number of parameters to be adjusted in the model (Table 5). The information criterias, based on the likelihood function, helps to select models with better quality of fit to the data. Higher information criteria values may be linked to the difficulty of adjustment of the parameterized models.

For the Wilmink function (Table 9), despite the higher estimated residual variance estimates are found for the smaller -2log(L), AIC and BIC criteria. Thus, the Wilmink function, despite its easier adjustment given the smaller number of parameters, is the one with better quality adjustment to the milk production data.

Similar results related to the quality of fit to the milk record data and to the simplicity to the computational adjustment of the Wilmink function were also reported by Olori et al. (1999), Macciotta, Vicario, and Cappio-Borlino (2005), Silvestre, Petim-Batista, and Colaco (2006), Strucken et al. (2011), and Dornelles et al. (2016).

The values of the estimates of the variance components for the additive genetic (VGA), animal permanent environmental (VEP) and phenotypic effects (VP) are similar between the fitted function and tended to increase during lactation (Table 6 and Figure 2). The highest estimates of VGA and VPE were observed in the final third of the lactation curve (Table 6 and Figure 2).

The variance for the permanent environmental effect was the most relevant cause for variation since the beginning of lactation (Table 6 and Figure 2). This effect, in the case of primiparous animals, may be associated to problems that occurred during growth or during the pre-pubertal period.

In a study by Schaeffer (2011), it was observed that the permanent environmental effects can be cumulative over the course of the productive life of the animal. According to the author, animals can suffer new environmental influences every day in their lives. So, it could have lasting environmental effects that influence milk production in the first milk record of a cow. During that first lactation, the cow would be influenced by new effects that would be added to those already present. This may explain the higher values of the permanent effect of variances observed at the end of lactation (Table 6 and Figure 2).

Notice that, by adjusting the LEG5 model, decreases in genetic and permanent environment variances in the last month of lactation were observed, which was not observed when adjusting the other functions (Table 6 and Figure 2). This may be associated with the number of model parameters (LEG5) that needed to be adjusted (Table 5) and the lower number of test day in the last milk recording observations (Table 3). In a study by Costa et al. (2008) and Dornelles et al. (2016), problems with the fit of random regressions to extremes of lactation data were also observed and related to the highest number of parameters of this function.

and ALI functions no changes were observed in the trends of the variance estimates at the beginning and at the end of lactation. These results of variance curve shapes follow the trend reported for Olori et al. (1999), Kettunen et al. (2000), and Naserkheil et al. (2016).

Low heritabilities were observed, considering the same test day in the adjustment of different functions. During lactation, the values of the heritability estimates for the different functions were similar (Table 7 and Figure 3).

The largest heritability values were observed in the final third of the lactation, apparently from 215 days (Table 7 and Figure 3). This indicates that, compared to the beginning of lactation, in the final third of lactation other differences between individual breeding values are more important in explaining the variation in the test day milk yield. This seems to be the most appropriate time applies selection in order to achieve the most genetic gain.

Table 7. Heritabilities for test day milk yields from five Holstein herds in first lactation created in the Minas Gerais State, Brazil, estimated by adjusting of different functions

DIM*	LEG3	LEG5	ALI	WIL
5	0.11	0.09	0.09	0.10
35	0.09	0.09	0.10	0.07
65	0.09	0.09	0.08	0.08
95	0.08	0.08	0.08	0.08
125	0.08	0.07	0.07	0.08
155	0.08	0.07	0.08	0.09
185	0.10	0.09	0.10	0.09
215	0.13	0.14	0.13	0.11
245	0.15	0.16	0.15	0.12
275	0.16	0.15	0.17	0.13
305	0.18	0.16	0.16	0.14

Note. *DIM = ays in milk; LEG3 = Legendre polynomials degree 3, order 4; LEG5 = Legendre polynomials degree 5, order 6; ALI = Ali & Schaeffer; WIL = Wilmink.
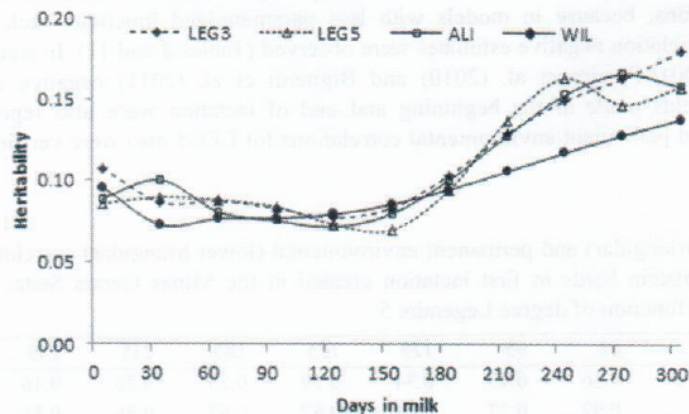

Figure 3. Heritabilities curves for test day milk yields for primiparous cows from five Holstein herds in Minas Gerais, Brazil, estimated by adjusting the functions of Legendre polynomials degree 3, order 4 (LEG3); Legendre polynomials degree 5, order 6 (LEG5); Ali & Schaeffer (ALI) and Wilmink (WIL)

The use of a covariance structure, as in the present study, to describe the permanent environment effect provides a reasonable strategy in separating the additive genetic effects from the permanent environmental effects on milk yield. Such strategy has also been reported by Kettunen, Mantysaari, and Poso (2000). Therefore, there is a reduction in the additive genetic variance, resulting also in reductions in estimates of heritability (Table 7 and Figure 3).

Table 10. Genetic (upper triangular) and permanent environmental (lower triangular) correlations for the test day milk yield from five Holstein herds in first lactation created in the Minas Gerais State, Brazil, obtained by adjusting the polynomial function of Ali & Schaeffer

DIM	5	35	65	95	125	155	185	215	245	275	305
5		0.44	0.46	0.36	0.22	0.05	-0.07	-0.15	-0.18	-0.19	-0.17
35	0.70		0.85	0.75	0.73	0.69	0.64	0.60	0.58	0.56	0.53
65	0.59	0.92		0.97	0.89	0.74	0.58	0.48	0.44	0.44	0.46
95	0.52	0.80	0.96		0.96	0.82	0.66	0.55	0.50	0.49	0.50
125	0.46	0.67	0.85	0.96		0.95	0.84	0.75	0.69	0.66	0.63
155	0.41	0.54	0.72	0.87	0.97		0.97	0.92	0.87	0.82	0.76
185	0.36	0.44	0.60	0.77	0.91	0.98		0.99	0.96	0.91	0.83
215	0.34	0.38	0.52	0.69	0.84	0.94	0.98		0.99	0.96	0.88
245	0.32	0.34	0.47	0.62	0.76	0.85	0.92	0.97		0.99	0.93
275	0.29	0.31	0.43	0.54	0.63	0.70	0.77	0.86	0.95		0.98
305	0.25	0.28	0.39	0.43	0.45	0.48	0.54	0.65	0.80	0.94	

Note. *DIM = Days in milk.

Table 11. Genetic (upper triangular) and permanent environmental (lower triangular) correlations for the test day milk yield from five Holstein herds in first lactation created in the Minas Gerais State, Brazil, obtained by adjusting the polynomial function of Wilmink

DIM	5	35	65	95	125	155	185	215	245	275	305
5		0.57	0.41	0.37	0.35	0.32	0.28	0.23	0.20	0.16	0.14
35	0.71		0.97	0.93	0.84	0.73	0.60	0.47	0.36	0.27	0.19
65	0.60	0.98		0.98	0.92	0.83	0.71	0.59	0.49	0.40	0.32
95	0.55	0.95	0.99		0.98	0.92	0.83	0.73	0.64	0.56	0.49
125	0.50	0.88	0.94	0.98		0.98	0.93	0.86	0.79	0.72	0.66
155	0.44	0.79	0.87	0.93	0.98		0.98	0.94	0.90	0.85	0.80
185	0.36	0.67	0.76	0.86	0.93	0.98		0.98	0.94	0.93	0.89
215	0.28	0.54	0.65	0.76	0.86	0.94	0.99		0.99	0.97	0.95
245	0.21	0.42	0.53	0.66	0.78	0.88	0.95	0.99		0.99	0.98
275	0.14	0.31	0.43	0.57	0.70	0.82	0.91	0.97	0.99		1.00
305	0.08	0.21	0.34	0.48	0.63	0.76	0.86	0.94	0.98	1.00	

Note. *DIM = Days in milk.

It was also noted that, from the 155th day of lactation, genetic and permanent environmental correlations between subsequent test days were of high magnitude, regardless of the fitted function (Tables 8, 9, 10 and 11). Possibly, genes that act on milk production in controls taken from the 155th day of lactation are the same. Breeding values for milk production in dairy cattle are highly related to genetic values of the milk yields in subsequent lactation controls. This implies that animals can be selected from this stage of lactation and that milk yields from not yet performed controls can be projected for the current lactation. This result is consistent with the observations already discusseded in this study related to the increased heritability observed at the end of the lactation. As a result, there will be a decrease on the generation interval and an increase on the number of daughters evaluated from each sire. Additionally, economic gains for producers can be observed, because lower productivity cows may be culled or replaced before completing their lactations. Correlations of high magnitude between consecutive milk weights taken from the 6th milk recording, close to 183 days of lactation, were also verified by Melo et al. (2005).

4. Conclusions

The Wilmink function is the most suitable for the study of primiparous cow test day milk yields from Holstein herds in the state of Minas Gerais, Brazil.

Hammami, H., Rekik, B., Soyeurt, H., Ben Gara, A., & Gengler, N. (2008). Genetic parameters for Tunisian Holsteins using a test-day random regression model. *Journal of Dairy Science, 91*(5), 2118-2126. https://doi.org/10.3168/jds.2007-0382

International Committee for Animal Recording. (2012). *International Agreement of Recording Practices.* Guidelines approved by the General Assembly held in Cork, Ireland on June 2012. Cork: ICAR.

Kettunen, A., Mantysaari, E. A., & Poso, J. (2000). Estimation of genetic parameters for daily milk yield of primiparous Ayrshire cows by random regression test-day models. *Livestock Production Science, 66*(3), 251-261. https://doi.org/10.1016/S0301-6226(00)00166-4

Kirkpatrick, M., Lofsvold, D., & Bulmer, M. (1990). Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics, 124*(4), 979-993.

Lethen, J. (1996). *Chebychev and empirical rules.* Texas: A&M University.

López-Romero, P., & Carabaño, M. J. (2000). Evaluación de modelos alternativos para el análisis de datos de producción de leche a lo largo de la lactación. *JTEA, 96A*(3), 318-336.

Macciotta, N. P. P., Vicario, D., & Cappio-Borlino, A. (2005). Detection of different shapes of lactation curve for milk yield in dairy cattle by empirical mathematical models. *Journal of Dairy Science, 88*(3), 1178-1191. https://doi.org/10.3168/jds.S0022-0302(05)72784-3

Mazerolle, M. J. (2004). *Mouvements et reproduction des amphibiens en tourbiéres perturbées.* Québec: Université Laval.

Melo, C. M. R., Packer, I. U., Costa, C. N., & Machado, P. F. (2005). Parâmetros genéticos para as produções de leite no dia do controle e da primeira lactação de vacas da raça Holandesa. *Revista Brasileira de Zootecnia, 34*(3), 796-806. https://doi.org/10.1590/S1516-35982005000300011

Meyer, K. (2001). Estimates of direct and maternal covariance functions for growth of Australian beef calves from birth to weaning. *Genetics Selection Evolution, 33*, 487-514. https://doi.org/10.1186/1297-9686-33-5-487

Misztal, I., Tsuruta, S., Strabel, T., Auvray, B., Druet, T., & Lee, D. H. (2002). BLUPF90 and related programs (BGF90). *Proceedings of 7th World Congress on Genetics Applied to Livestock Production* (pp. 19-23). Montpelier: Inra.

Naserkheil, M., Miraie-Ashtiani, S. R., Nejati-Javaremi, A., Son, J., & Lee, D. (2016). Random Regression Models Using Legendre Polynomials to Estimate Genetic Parameters for Test-day Milk Protein Yields in Iranian Holstein Dairy Cattle. *Asian-Australasian Journal of Animal Sciences, 29*(12), 1682-1687. http://doi.org/10.5713/ajas.15.0768

Olori, V. E., Brotherstone, S., Hill, W. G., & McGuirk, B. J. (1999). Fit of standard models of the lactation curve to weekly records of milk production of cows in a single herd. *Livestock Production Science, 58*(1), 55-63. https://doi.org/10.1016/S0301-6226(98)00194-8

Pereira, R. J., Lopes, P. S., Verneque, R. S., Santana Júnior, M. L., Lagrotta, M. R., & Torres, R. A. (2010). Funções de covariância para produção de leite no dia do controle em bovinos Gir leiteiro. *Pesquisa Agropecuária Brasileira, 45*(11), 1303-1311. https://doi.org/10.1590/S0100-204X2010001100011

Poll, H. (2013). *Campo bom.* Anuário Brasileiro da Pecuária, Santa Cruz do Sul: Editora Gazeta Santa Cruz.

R Core Team. (2013). *R: A language and environment for statistical computing.* Vienna: R Foundation for Statistical Computing.

Rebouças, G. F., Gonçalves, T. M., Martinez, M. L., Azevedo Junior, J., & Koops, W. (2008). Novas funções para estimar a produção de leite, em 305 dias de lactação, de vacas da raça Gir. *Revista Brasileira de Zootecnia, 37*(7), 1222-1229. https://doi.org/10.1590/S1516-35982008000700012

Schaeffer, L. R. (2011), Cumulative permanent environmental effects for repeated records animal models. *Journal of Animal Breeding and Genetics, 128*, 95-99. https://doi.org/10.1111/j.1439-0388.2010.00894.x

Schaeffer, L. R., Jamrozik, J., Kistemaker, G. J., & Van Doormaal, J. (2000). Experience with a test-day model. *Journal of Dairy Science, 83*(5), 1135-1144. https://doi.org/10.3168/jds.S0022-0302(00)74979-4

Silvestre, A. M., Petim-Batista, F., & Colaço, J. (2006). The accuracy of seven mathematical functions in modeling dairy cattle lactation curves based on test-day records from varying sample schemes. *Journal of Dairy Science, 89*(5), 1813-1821. https://doi.org/10.3168/jds.S0022-0302(06)72250-0