

Comparison of random regression models to estimate genetic parameters for milk production in Guzerat (*Bos indicus*) cows

D.J.A. Santos¹, M.G.C.D. Peixoto², R.R. Aspilcueta Borquis¹,
R.S. Verneque², J.C.C. Panetto² and H. Tonhati¹

¹Departamento de Zootecnia, Faculdade de Ciências Agrárias e Veterinárias,
Universidade Estadual Paulista “Júlio de Mesquita Filho”, Jaboticabal, SP, Brasil
²Embrapa Gado de Leite, Juiz de Fora, MG, Brasil

Corresponding author: D.J.A. Santos
E-mail: daniel_jordan2008@hotmail.com

Genet. Mol. Res. 12 (1): 143-153 (2013)
Received March 6, 2012
Accepted November 7, 2012
Published January 24, 2013
DOI <http://dx.doi.org/10.4238/2013.January.24.6>

ABSTRACT. Random regression models have been widely used to estimate genetic parameters that influence milk production in *Bos taurus* breeds, and more recently in *B. indicus* breeds. With the aim of finding appropriate random regression model to analyze milk yield, different parametric functions were compared, applied to 20,524 test-day milk yield records of 2816 first-lactation Guzerat (*B. indicus*) cows in Brazilian herds. The records were analyzed by random regression models whose random effects were additive genetic, permanent environmental and residual, and whose fixed effects were contemporary group, the covariable cow age at calving (linear and quadratic effects), and the herd lactation curve. The additive genetic and permanent environmental effects were modeled by the Wilmink function, a modified Wilmink function (with the second term divided by 100), a function that combined third-order Legendre polynomials with the last term of the Wilmink function, and the Ali and Schaeffer function. The residual variances were modeled by means of 1, 4, 6, or 10 heterogeneous classes, with the exception of the last term of the Wilmink function, for which there were 1,

3, 6, or 10 classes. The models gave similar heritability estimates, ranging from 0.20 to 0.33. Genetic correlations between adjacent records were high values (0.83-0.99), but they declined when the interval between the test-day records increased, and were negative between the first and last records. The model employing the Ali and Schaeffer function with six residual variance classes was the most suitable for fitting the data.

Key words: Covariance function; Parametric functions; Test-day milk yield; Zebu

INTRODUCTION

Random regression models (RRMs) have been widely used for analysis of traits whose measures are obtained sequentially over time (longitudinal data), such as milk production. According to Meyer and Hill (1997), RRM are equivalent to covariance functions, described initially by Kirkpatrick et al. (1990). Covariance functions permit description of the changes in covariance between measures that occur over time, and can predict the variances for points of the trajectory with little or no information. They are obtained in RRM by means of the co(variance) matrix between the random regression coefficients (Meyer, 1999).

There are different functions that can be used to fit the yield trajectory. Among those used in random regression, the ones most often used are parametric functions and Legendre orthogonal polynomials. Legendre polynomials are flexible, enabling fitting curves independently of the trait of interest. In turn, parametric functions are based on components of the typical curve and tend to impose a particular shape, which can result in satisfactory fits when the data take this form but unsatisfactory fits when the trait's data follow a different path.

Among the parametric functions used in RRM to study test-day milk yield (TMDY), the exponential function of Wilmink and the logarithmic one of Ali and Schaeffer are more prominent (López-Romero and Carabaño, 2003; de Melo et al., 2007; Herrera et al., 2008; Pereira et al., 2010). Modifications of the Wilmink function have been proposed with the intention of improving the adjustment of curves to the data. In this respect, Jakcbozen et al. (2002) proposed dividing the second term by 100 to reduce the amplitude of the covariable, with the aim of improving the model's numerical properties. Brotherstone et al. (2000) altered the value of the parameter a_3 , comparing the standard value (-0.05) with -0.068 and -0.10, and concluded that the value -0.068 provided the best fit for Dutch breeds. They also obtained a higher value for the maximum log-likelihood and lower incidence of negative correlations between the initial and final test-day records. Freitas (2003), who studied the Gyrolando breed, tested the same values and concluded that -0.05 was the best. Lindauer and Mäntysaari (1999) proposed a function formed by a combination of the Wilmink and third-order Legendre orthogonal polynomials for better description of the lactation curve.

For an adequate partition of the total variance, studies with RRM have concluded that heterogeneous residual variances must be considered (El Faro and Albuquerque, 2003; Bignardi et al., 2009; Takma and Akbas, 2009). RRM that consider heterogeneous residual variances are more suitable for fitting the data than are models that consider the residual variance as being homogeneous, since these tend to overestimate the additive variance (Jamrozik and Schaeffer, 1997).

The aim of this study was to compare RRM with different residual variance structures to ana-

lyze milk yield from first lactations of Guzerat cows by using the parametric functions of Wilmink and Ali and Schaeffer and a combination of third-order Legendre polynomials and the Wilmink function.

MATERIAL AND METHODS

The data used consisted of the TDMY records from the first lactation of Guzerat cows (Table 1), with calving recorded between 1987 and 2009 and ages between 23 and 65 months. The average milk production for the TDMY records was 6.72 ± 2.45 kg. The pedigree file contained 10,753 animals. The data were obtained from the National Breeding Program for Guzerat Dairy Cattle Improvement, coordinated by the Embrapa Dairy Cattle Research Unit (Embrapa Dado de Leite) in partnership with the Brazilian Center for Guzerat Breeding and the Brazilian Zebu Breeders Association.

Table 1. Description of data used in this study.

Description of data	Number of records
Number of test-day records	20,524
Number of recorded animals	2,816
Number of sires	371
Number of dams	1,774
Number of herds	28
Number of contemporary groups	401

We considered TDMY records from the 6th to 305th days of lactation. The records were divided into 10 monthly classes, and only data from cows with at least 4 records were kept in the data set.

The contemporary groups were formed by herd, year, and test-day season, which was divided into the dry season (April to September) and the rainy season (October to March). Contemporary groups were maintained, provided they contained at least 3 animals. The data are described in Table 1.

The analyses were performed using a single-trait RRM. This model included the direct additive genetic, permanent environmental, and residual effects as the random effects and the contemporary group, average herd lactation curve, and linear and quadratic effects of the covariable cow age at calving as the fixed effects. The variance components were estimated using the restricted maximum likelihood method with the Wombat program (Meyer, 2006).

The lactation curve was fitted by the logarithmic function of Ali and Schaeffer (AS; 1987), the exponential function of Wilmink (Wl; 1987), a modified Wilmink function (Wlm) with the second term divided by 100 (Jakcosen et al., 2002), and a combination of the Wilmink function with third-order Legendre polynomials (LM; Lindauer and Mäntysaari, 1999). The mean herd curve and the random regressions for the additive genetic and permanent environmental effects were modeled according to the function used.

For the residual variances, both homogeneous and heterogeneous structures were considered, with different numbers of classes. The model with 10 classes considered each month as a different class, while the models with fewer classes were grouped according to the similarity of the variances. Therefore, the model adjusted by the AS considered 6 residual variance classes, grouped in the following form: 1, 2, 3-5, 6-7, 8-9, 10; while the grouping with 4 classes was 1, 2, 3-9, 10. The model adjusted by the Wl and Wlm considered 6 residual variance classes, grouped: as 1, 2-3, 4, 5-8, 9, 10, or 4 classes, grouped as 1, 2-8, 9, 10. The model adjusted by LM included 6 residual

variance classes, grouped 1-2, 3, 4-5, 6-7, 8-9, 10, or 3 classes, grouped as 1-2, 3-9, 10.

In the WI given by $y = a_0 + a_1t + a_2 \exp(a_3t)$, where t is the number of days in lactation and the parameter a_3 is related to the curve's shape and peak lactation moment. This was considered as a constant, to reduce the number of parameters to be estimated by the models. To estimate the value of the maximum log-likelihood function, the following values of a_3 were evaluated: -0.06; -0.05 (standard); -0.04; -0.03; -0.025; -0.02; -0.015, and -0.01.

The random regression model used can be represented mathematically as follows:

$$y = Xb + Za + Wap + e$$

where y is the vector of the N observations measured in N_a animals; b is the fixed effect vector, including the solutions for contemporary group and the covariable age at calving (linear and quadratic regressions); a is the vector of solutions for the coefficients of the additive genetic random effects; ap is the vector of solutions for the coefficients of the permanent environmental random effects; e is the vector of the N different residuals, and X , Z , W are incidence matrices for the fixed and direct random genetic and permanent environmental effects, respectively.

The assumptions about the components of the model are as follows:

$$E \begin{bmatrix} y \\ a \\ ap \\ e \end{bmatrix} = \begin{bmatrix} Xb \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \begin{matrix} V(a) = K_A \otimes A; \\ e \quad V(ap) = K_{AP} \otimes I_{Nd}; \\ V(e) = R, \end{matrix}$$

where K_A and K_{AP} are the co(variance) matrices between the random regression coefficients of the additive genetic and permanent environmental effects, respectively; A is the pedigree matrix among the individuals; I_{Nd} is an identity matrix; R represents a matrix containing the residual variances.

The models' performances were compared using the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) of Schwarz (1978), according to which lower values indicate a better model (Wolfinger, 1993).

The information criteria can be represented as follows:

$$\begin{aligned} \text{AIC} &= -2\log L + 2p, \\ \text{BIC} &= -2\log L + p \log(N - r) \end{aligned}$$

where p is the number of model parameters, N is the total number of records, r is the rank of the fixed effect incidence matrix in the model, and $\log L$ is the natural logarithm of the likelihood function.

RESULTS AND DISCUSSION

Table 2 contains the results of the maximum log-likelihood function for the models that considered homogeneous variances, fitted by the WI with different values of the parameter a_3 . The best value of this parameter among these models was -0.025, because it obtained the highest log-likelihood value. A similar result was found by Pereira et al. (2010), who evaluated working with data on the Gyr breed. In WI, the parameter a_3 is related to

the peak milk production, and the standard value of this parameter is -0.05, obtained from fitting lactation curves mainly of *Bos taurus* breeds (Wilmink, 1987), for which this peak is frequently reached on the 60th lactation day. In Zebu breeds, however, the peak milk yield occurs closer to the start of lactation, with a short ascension period to the peak, probably explaining the better fit with a lower value for this parameter (Cobuci et al., 2000).

Table 3 shows the values of the log-likelihood function (logL), the AIC, and BIC.

Table 2. Value of the maximum log-likelihood function (logL) for the models considering homogeneous variances, fitted by the Wilmink exponential function (WI) with different values of the parameter a_3 .

a_3	logL
-0.06	-16.323
-0.05	-16.308
-0.04	-16.291
-0.03	-16.276
-0.025	-16.273
-0.02	-16.276
-0.015	-16.292
-0.01	-16.327

Value in bold indicates the best model based on logL.

Table 3. Number of classes of residual variance (e), number of parameters (P), value of log likelihood function (logL), Akaike's information criterion (AIC) and Schwarz's Bayesian information criterion (BIC) for the random regression models using parametric functions Wilmink, Wilmink modified by the combination of the Legendre polynomials with the Wilmink function and for Ali and Schaeffer function.

Model	e	P	logL	AIC	BIC
Wilmink					
WI-1	1	13	-16.273	32.572	32.675
WI-4	4	16	-16.264	32.561	32.687
WI-6	6	18	-16.243	32.522	32.664
WI-10	10	22	-16.217	32.479	32.653
Wilmink modified					
Wlm-1	1	13	-16.268	32.563	32.666
Wlm-4	4	16	-16.260	32.552	32.678
Wlm-6	6	18	-16.238	32.513	32.655
Wlm-10	10	22	-16.213	32.470	32.644
Legendre + Wilmink					
LM-1	1	21	-16.023	32.089	32.256
LM-3	3	23	-15.999	32.045	32.227
LM-6	6	26	-15.976	32.005	32.211
LM-10	10	30	-15.975	32.011	32.249
Ali and Schaeffer					
AS-1	1	31	-15.925	31.912	32.157
AS-4	4	34	-15.916	31.901	32.170
AS-6	6	36	-15.900	31.873	32.157
AS-10	10	40	-15.899	31.879	32.196

Values in bold indicate the lowest values for AIC and BIC.

In the present study, the logL values improved with an increasing number of parameters. The values of the AIC and BIC tests to compare the models showed that the models considering homogeneous residual variances fit the data worse than those considering heterogeneous variances, irrespective of the function employed. However, two of the models adjusted by the Ali and Schaeffer function (AS-1 and AS-6), which considered homogeneous

variances, presented the best BIC values, and the second of these obtained better logL and AIC values than the first. These results thus indicate that the residual variance behaves differently during lactation, as also observed by Costa et al. (2005), Herrera et al. (2008), and Pereira et al. (2010), studying parametric functions for milk production by the Zebu Gyr breed.

The Wilmink function obtained worse values for logL, AIC, and BIC than the modified Wilmink function, both with the same number of parameters. The best results for the Wilmink function were obtained by the modified function with 10 residual variance classes (Wlm-10), according to both AIC and BIC. The best model combining LM with the Wilmink function was that with 6 residual variance classes. For AS the model with 6 residual variance classes (AS-6) obtained the lowest values for both AIC and BIC and was therefore considered the best model. BIC is more rigorous than AIC and tends to indicate more parsimonious models, so we used the model that considered homogeneous variances (AS-1) in the subsequent analyses because it had the same BIC value as model AS-6 but a smaller number of parameters.

The models fitted by a combination of LM and the Wilmink function obtained better results than those using only the Wilmink function, in either its original or modified specification. AS was thus better than the other parametric functions, according to the criteria evaluated, and was consistent with the results of most studies conducted with other breeds in Brazil (de Melo et al., 2007; Herrera et al., 2008; Pereira et al., 2010).

The models fitted by AS function obtained the highest magnitudes for the covariances and random regression coefficients (Table 4). The covariances of the modified Wilmink function involving the parameter a_i of the additive genetic and permanent environmental coefficients were 100 times larger than those of the WI10 function (data not shown), and the variance was 10,000 times larger, making the values of the covariance matrix more regular without altering the correlations between these coefficients. However, the modification of the Wilmink

Table 4. Estimates of variances (diagonal), covariance (above the diagonal) and correlations (below the diagonal) between random regression coefficients and eigenvalues (λ) of the coefficients matrix for additive genetic effect and permanent environment effect for the best models of each function.

	Additive genetic effect					Permanent environment effect				
	α_0	α_1	α_2	α_3	α_4	α_0	α_1	α_2	α_3	α_4
Wlm-10										
α_0	3.18	-1.13	-2.36	-	-	3.92	-0.98	-2.68	-	-
α_1	-0.88	0.51	0.75	-	-	-0.73	0.46	0.92	-	-
α_2	-0.67	0.53	3.83	-	-	-0.57	0.58	5.54	-	-
λ	6.19	1.24	0.09	-	-	7.88	1.80	0.19	-	-
LM-6										
α_0	2.02	-0.70	0.10	-2.01	-	3.79	-0.77	0.25	-2.55	-
α_1	-0.57	0.74	-0.03	0.80	-	-0.35	1.24	-0.63	3.57	-
α_2	0.27	-0.13	0.07	-0.06	-	0.15	-0.64	0.79	-2.80	-
α_3	-0.83	0.55	-0.14	2.89	-	-0.34	0.82	-0.80	15.3	-
λ	4.79	0.48	0.38	0.06	-	17.17	3.29	0.38	0.25	-
AS-6										
α_0	93.56	-131.95	41.71	-56.41	9.27	1148.6	-1787.4	656.85	-673.91	104.52
α_1	-0.99	189.81	-62.74	79.76	-13.24	-0.99	2820.60	-1063.40	1042.70	-160.87
α_2	0.88	-0.93	23.87	-25.26	4.26	0.94	-0.97	421.66	-377.98	57.66
α_3	-0.98	0.97	-0.871	35.27	-5.87	-0.97	0.98	-0.92	398.42	-62.06
α_4	0.96	-0.97	0.877	-0.99	0.98	0.99	-0.97	0.90	-0.99	9.69
λ	337.25	4.88	1.07	0.30	0.00	4751.92	45.10	1.27	0.71	0.00

Values in bold indicate the eigenvalues in which to base the discussion of the results showed in this table.

function increased the explanatory power of the third eigenvalue from 0 to 1.25 and 1.88% for the additive genetic and permanent environmental effects, respectively. The last eigenvalue of the models fitted by the AS explained 0% of the total variation in the additive genetic and permanent environmental effects. This indicates that these models were overparameterized, despite adequately fitting the data.

The first eigenvalue of the covariance matrix between the regression coefficients explained a higher proportion of the total variation in the models adjusted by the AS (more than 96%), both for the genetic effect and the permanent environmental effect. For the other models, the first eigenvalue was responsible for up to 78% of the total variation.

The phenotypic and additive genetic variances estimated by the Wlm-10, LM-6, AS-1, and AS-6 models showed similar tendencies and coincided at many points (Figure 1). There was no coincidence of points when using the Wlm-10 model only for the last 2 months, with the difficulty of fitting the data attributed to the end of lactation. With respect to the permanent environmental variance, the tendencies among the models were also similar. However, estimate of the permanent environmental variance for the first lactation month obtained with the AS-6 model was lower than those obtained with the other models.

The residual variances presented a similar and constant tendency from the 3rd to the 9th lactation month. López-Romero and Carabaño (2003) reported that for the interval between 75 and 275 days in milk, the residual variances can be assumed to be homogeneous. This can explain the adequate fit produced by model AS-1, which considered constant variance, and the value near those obtained by the other models for this interval. In model AS-6, the residual variance estimates were greater in the first month and smaller in the second and last months than those of the other models, this being the model that obtained the largest magnitude of differences between the estimates. The Wlm-10 and LM-6 models tended to produce the highest estimates for the second month, but diverged in the last month of lactation, where model Wlm-10 presented a higher residual variance estimate than LM-6.

The heritability estimates produced by models Wlm-10, LM-6, AS-1, and AS-6 behaved similarly (Figure 2), except for the second and last months, where model Wlm-10 produced smaller estimates than the others. The amplitudes of variation in the heritability estimates of the models ranged from 0.20 to 0.33, with the lowest heritability estimates obtained for TDMY in the eighth month and the highest in the second or third month. Freitas et al. (2010), who also evaluated data on the Guzerat breed, found similar results but observed the highest heritability values at the start of lactation. Heritability estimates similar to those in the present study, obtained using Wilmink and AS functions, were reported by Herrera et al. (2008) and Pereira et al. (2010), with values varying from 0.15 to 0.33 for TDMY of Gyr cows. Araújo et al. (2006), who studied the Dutch breed, also found heritability estimates with this amplitude by using the same functions.

The estimates of the genetic and phenotypic correlations of the records, obtained by models Wlm-10, LM-6, and AS-6, are presented in Tables 5, 6, and 7, respectively. The genetic correlations were high (near 1) between the yields of adjacent records and declined as the interval between the increased records. This result is similar to those found by Biguardi et al. (2009) for the Dutch breed and by Kettunen et al. (2000) for the Ayrshire breed; RRM models were used in both studies.

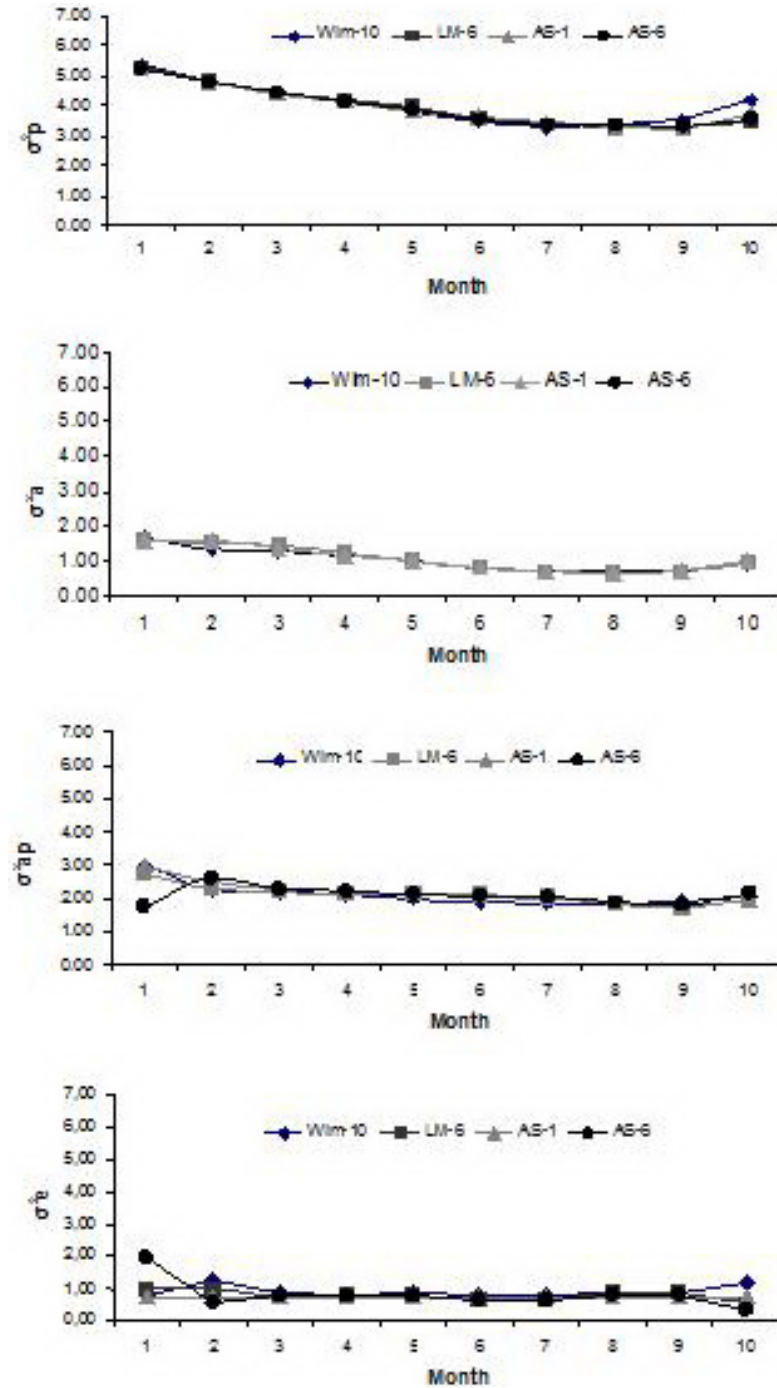


Figure 1. Phenotypic (σ^2_p), genetic (σ^2_a), permanent environmental (σ^2_{ep}), and residual (σ^2_e) variances estimated for monthly milk yield obtained with the Wlm-10, LM-6, AS-1, and AS-6 random regression models.

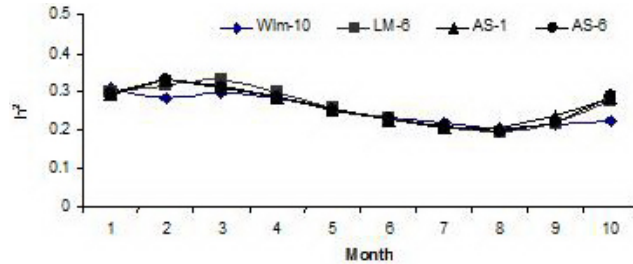


Figure 2. Heritability estimates (h^2) for monthly milk yield obtained with the Wlm-10, LM-6, AS-1, and AS-6 random regression models.

Table 5. Phenotypic (below diagonal) and genetic (above diagonal) correlation estimates between monthly milk yield obtained with the Wlm-10 random regression model.

Month	1	2	3	4	5	6	7	8	9	10
1	-	0.83	0.68	0.57	0.49	0.39	0.27	0.13	-0.02	-0.15
2	0.68	-	0.97	0.92	0.86	0.77	0.63	0.44	0.23	0.02
3	0.59	0.76	-	0.99	0.95	0.88	0.76	0.58	0.37	0.14
4	0.52	0.72	0.80	-	0.99	0.94	0.84	0.68	0.48	0.27
5	0.48	0.68	0.77	0.78	-	0.98	0.91	0.79	0.61	0.41
6	0.44	0.64	0.73	0.76	0.77	-	0.97	0.89	0.75	0.58
7	0.39	0.58	0.67	0.71	0.74	0.76	-	0.97	0.88	0.75
8	0.33	0.49	0.58	0.63	0.67	0.72	0.75	-	0.97	0.89
9	0.27	0.40	0.48	0.54	0.60	0.66	0.72	0.74	-	0.97
10	0.20	0.28	0.35	0.42	0.49	0.57	0.64	0.69	0.73	-

Table 6. Phenotypic (below diagonal) and genetic (above diagonal) correlation estimates between monthly milk yield obtained with the LM-6 random regression model.

Month	1	2	3	4	5	6	7	8	9	10
1	-	0.87	0.73	0.62	0.51	0.39	0.26	0.13	0.01	-0.06
2	0.68	-	0.97	0.92	0.85	0.76	0.63	0.47	0.3	0.13
3	0.61	0.79	-	0.99	0.95	0.88	0.77	0.62	0.42	0.22
4	0.54	0.73	0.8	-	0.99	0.94	0.86	0.71	0.52	0.3
5	0.48	0.66	0.75	0.78	-	0.98	0.92	0.8	0.62	0.4
6	0.44	0.59	0.7	0.76	0.8	-	0.98	0.89	0.73	0.52
7	0.4	0.53	0.63	0.71	0.76	0.81	-	0.97	0.86	0.68
8	0.35	0.45	0.55	0.62	0.68	0.74	0.76	-	0.96	0.84
9	0.3	0.4	0.48	0.53	0.58	0.64	0.69	0.72	-	0.96
10	0.25	0.34	0.38	0.4	0.43	0.48	0.55	0.63	0.74	-

Table 7. Phenotypic (below diagonal) and genetic (above diagonal) correlation estimates between monthly milk yield obtained with the AS-6 random regression model.

Month	1	2	3	4	5	6	7	8	9	10
1	-	0.84	0.74	0.62	0.51	0.38	0.26	0.13	0.03	-0.03
2	0.69	-	0.97	0.9	0.83	0.74	0.62	0.47	0.28	0.1
3	0.61	0.8	-	0.98	0.94	0.88	0.77	0.62	0.42	0.21
4	0.54	0.72	0.8	-	0.99	0.95	0.86	0.72	0.52	0.3
5	0.47	0.65	0.75	0.79	-	0.99	0.93	0.81	0.62	0.4
6	0.43	0.59	0.69	0.76	0.8	-	0.98	0.89	0.74	0.53
7	0.38	0.53	0.62	0.69	0.75	0.8	-	0.97	0.86	0.68
8	0.32	0.47	0.53	0.6	0.67	0.73	0.77	-	0.96	0.84
9	0.28	0.4	0.46	0.51	0.57	0.64	0.69	0.73	-	0.96
10	0.22	0.32	0.38	0.41	0.43	0.48	0.55	0.63	0.76	-

All the models presented negative correlations between milk production in the first and last lactation months, with model Wlm-10 producing the lowest correlation and also a negative correlation between the first and ninth month. Negative estimates for genetic correlations by using RRMJs adjusted by parametric functions were also found by Brotherstone et al. (2000), López-Romero and Carabaño (2003), and de Melo et al. (2007) for the Dutch breed; by Kettunen et al. (2000) for the Ayrshire breed, and by Costa et al. (2005) and Pereira et al. (2010) for the Gyr breed. However, Cobuci et al. (2005), who used the Wilmink function, and Herrera et al. (2008), who used the Wilmink and AS functions to fit the additive genetic and permanent environmental effects, did not observe a negative correlation with data for the Dutch and Gyr breeds, respectively. The negative correlations obtained in this study were possibly attributable to the difficulty of using RRMJs to model the start and end of lactation, since there are fewer records in these periods, particularly at the end of lactation, because of the high frequency of short lactation periods in the Guzerat breed.

For model AS-1 (data not shown), the estimates of the genetic correlations were slightly lower than those produced by model AS-6, while the estimates of the phenotypic correlations were slightly higher. Models Wlm-10 and Wl-10 did not present relevant differences between the genetic and phenotypic correlations.

According to the statistical criteria and estimates of the genetic parameters considered (mainly genetic correlation), the Wilmink function produced the worst fits, making it less recommended for describing the milk yield curve for the Guzerat breed. The combination of the LM and the Wilmink function produced genetic parameter estimates similar to those of the other models, but with a smaller number of parameters than model AS-6. Therefore, the combination of the LM and the Wilmink function can be used to fit the lactation curve for genetic evaluation of the breed, despite having produced an inferior result according to the statistical criteria adopted (AIC and BIC). The models adjusted by AS produced the best fits, with the model with 6 residual variance classes being the best. The model fitted by the AS with homogeneous residual variance presented estimates of the genetic parameters and residual variances near those obtained by the best model (AS-6) during most of the lactation period, so it can be used as an alternative, more parsimonious, model for fitting the data.

CONCLUSIONS

The results obtained by this study indicate the need to consider heterogeneous residual variances when using the Wilmink functions and the function that combines LM with the Wilmink function. AS produced the best fit for the functions evaluated, and the models adjusted by it were adequate for the data in the present study, irrespective of the residual variance structure.

ACKNOWLEDGMENTS

Research supported by the National Council of Technological and Scientific Development (CNPq). The authors thank Associação Brasileira dos Criadores de Guzerá (ACBG) and FAPEMIG.

REFERENCES

- Ali TE and Schaeffer LR (1987). Accounting for covariances among test day milk yields in dairy cows. *Can. J. Anim. Sci.* 67: 637-644.

- Araújo CV, Torres RA, Costa CN, Torres Filho RA, et al. (2006). Uso de modelos de regressão aleatória para descrever a variação genética da produção de leite na raça Holandesa. *Rev. Bras. Zootec.* 35: 975-981.
- Bignardi AB, El Faro L, Cardoso VL, Machado PF, et al. (2009). Random regression models to estimate test-day milk yield genetic parameters Holstein cows in Southeastern Brazil. *Livest. Prod. Sci.* 123: 1-7.
- Brotherstone S, White IMS and Meyer K (2000). Genetic modeling of daily yield using orthogonal polynomials and parametric curves. *Anim. Sci.* 70: 407-415.
- Cobuci JA, Euclides RF, Verneque RS, Teodoro RL, et al. (2000). Curva de lactação na raça Guzerá. *Rev. Bras. Zootec.* 29: 1332-1339.
- Cobuci JA, Euclides RF, Lopes PS, Costa CN, et al. (2005). Estimation of genetic parameters for test-day milk yield in Holstein cows using a random regression model. *Genet. Mol. Biol.* 28: 75-83.
- Costa CN, Melo CMR, Machado CHC, Freitas AF, et al. (2005). Parâmetros genéticos para a produção de leite de controles individuais de vacas da raça Gir Leiteiro estimados com modelos de repetibilidade e regressão aleatória. *Rev. Bras. Zootec.* 34: 1519-1530.
- de Melo CM, Packer IU, Costa CN and Machado PF (2007). Genetic parameters for test day milk yields of first lactation Holstein cows by random regression models. *Animal* 1: 325-334.
- El Faro L and Albuquerque LG (2003). Utilização de modelos de regressão aleatória para produção de leite no dia do controle, com diferentes estruturas de variâncias residuais. *Rev. Bras. Zootec.* 32: 1104-1113.
- Freitas LS, Silva MA, Verneque RS, Valente BD, et al. (2010). Avaliação da persistência na lactação da raça Guzerá, utilizando modelos de regressão aleatória. *Arq. Bras. Med. Vet. Zootec.* 62: 401-408.
- Freitas MS (2003). Utilização de Modelos de Regressão Aleatória na Avaliação Genética de Animais da Raça Girolando. Master's thesis, Universidade Federal de Viçosa, Viçosa.
- Herrera LGG, El Faro L, Albuquerque LG and Tonhati H et al. (2008). Estimativas de parâmetros genéticos para produção de leite e persistência da lactação em vacas Gir, aplicando modelos de regressão aleatória. *Rev. Bras. Zootec.* 37: 1584-1594.
- Jakobsen JH, Madsen P, Jensen J, Pedersen J, et al. (2002). Genetic parameters for milk production and persistency for Danish Holsteins estimated in random regression models using REML. *J. Dairy Sci.* 85: 1607-1616.
- Jamrozik J and Schaeffer LR (1997). Estimates of genetic parameters for a test day model with random regressions for yield traits of first lactation Holsteins. *J. Dairy Sci.* 80: 762-770.
- Kettunen A, Mäntysaari EA and Pösö J (2000). Estimation of genetic parameters for daily milk yield of primiparous Ayrshire cows by random regression test-day models. *Livest. Prod. Sci.* 66: 251-261.
- Kirkpatrick M, Lofsold D and Bulmer M (1990). Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics* 124: 979-993.
- Lidauer M and Mäntysaari EA (1999). Multiple trait reduced rank random regression test-day model for production traits. *Interbull Bull.* 22: 74-80.
- López-Romero P and Carabaño MJ (2003). Comparing alternative random regression models to analyse first lactation daily milk yield data in Holstein-Friesian cattle. *Livest. Prod. Sci.* 82: 81-96.
- Meyer K (1999). Estimates of genetic and phenotypic covariance functions for postweaning growth and mature weight of beef cows. *J. Anim. Breed. Genet.* 116: 181-205.
- Meyer K (2006). WOMBAT - A Program for Mixed Model Analyses by Restricted Maximum Likelihood. User Notes. Animal Genetics and Breeding Unit, Amidale.
- Meyer K and Hill WG (1997). Estimation of genetic and phenotypic covariance functions for longitudinal or "repeated" records by restricted maximum likelihood. *Livest. Prod. Sci.* 47: 185-200.
- Pereira RJ, Lopes OS, Verneque RS, Santana Júnior ML, et al. (2010). Funções de covariância para produção de leite no dia do controle em bovinos Gir leiteiro. *Pesq. Agropec. Bras.* 45: 1303-1311.
- Schwarz G (1978). Estimating the dimension of a model. *Ann. Statist.* 6: 461-464.
- Takma C and Akabas Y (2009). Heterogeneity of residual variances of test day milk yields estimated by random regression model in Turkish Holsteins. *J. Anim. Vet. Adv.* 8: 782-787.
- Wilmink JBM (1987). Adjustment of test-day milk, fat and protein yields for age, season and stage of lactation. *Livest. Prod. Sci.* 16: 335-348.
- Wolfinger R (1993). Covariance structure selection in general mixed models. *Commun. Stat.* 22: 1079-1106.