Genetic parameters of test-day milk yield in Guzerá cattle under tropical conditions

D.A.C. Cruz[1], M.G.C.D. Peixoto[2], F.A.T. Bruneli[2], A.B. Bignardi[3] and L. El Faro[1]

[1]Centro de Pesquisas de Bovinos de Corte, Instituto de Zootecnia, Sertãozinho, SP, Brasil
[2]Embrapa Gado de Leite, Juiz de Fora, MG, Brasil
[3]Grupo de Melhoramento Animal de Mato Grosso,
Instituto de Ciências Agrárias e Tecnológicas, Universidade Federal de Mato Grosso, Rondonópolis, MT, Brasil

Corresponding author: L. El Faro
E-mail: lenira@iz.sp.gov.br

ABSTRACT. The objective of this study was to estimate genetic parameters for test-day milk yield (TDMY) in Guzerá cows using random regression models. Additive and permanent environmental random effects were modeled by random regression on fourth- and fifth-order orthogonal Legendre polynomials, respectively. The residual variances were heterogeneous, with seven classes. Heritability estimates for TDMY ranged from 0.24 to 0.52, with higher heritabilities for yields during early lactation. Genetic correlations between TDMYs ranged from -0.03 to 0.95. The phenotypic and permanent environmental correlations were all positive, and the highest estimates were between adjacent TDMYs. The results suggest that TDMYs obtained with random regression models may be used as selection criteria for Guzerá cattle.

Key words: Lactation curve; Legendre polynomial; Longitudinal data

INTRODUCTION

In many countries, genetic evaluations of milk yield are based on cumulative production up to 305 days of lactation (P305). Formulas are used in the calculation of P305, and projection factors are used in cases of ongoing or short lactations. The main problem in using these factors is related to the fact that there is no variability in the shape of the lactation curves; consequently, some of the genetic variability of the traits is eliminated (Jakobsen et al., 2002).

Random regression models (RRMs) have been proposed as an alternative approach for the genetic evaluation of longitudinal data, such as test-day milk yield (TDMY). These models are also known as infinite dimensional models, because infinite measures exist between two points on a line and these measures are more correlated the closer they are (Meyer, 1998; Herrera et al., 2008; Bignardi et al., 2009a; Santos et al., 2013).

In analyzing data related to milk production, these models have the advantage of fitting TDMY by permitting the inclusion of specific environmental effects, thus reducing residual variance. Furthermore, these models permit the estimation of genetic lactation curves for each animal as a deviation from the average population curve, and when allowed by the data structure, the estimation of lactation curves for groups of animals (Bignardi et al., 2009b).

Among the Zebu breeds present in Brazil, Guzerá cattle are one of the most productive. This breed is considered rustic, and is widely used in strategic crossings with the aim of producing animals adapted to climatic conditions. In Brazil, Guzerá breeding program objectives are focused on dual-purpose selection (Peixoto et al., 2010); however, the breed is characterized by a high frequency of short lactations, and since it is not a specific dairy breed, by less persistent lactation curves. In view of the above considerations, the objective of the present study was to estimate genetic parameters for milk yield in Guzerá cows using RRMs.

MATERIAL AND METHODS

The data were provided by the National Breeding Program for the Improvement of Guzerá for Milk Purposes (Programa Nacional de Melhoramento do Guzerá para Leite), which is coordinated by Embrapa Gado de Leite in collaboration with the Brazilian Center for the Genetic Improvement of Guzerá Cattle (Centro Brasileiro de Melhoramento Genético do Guzerá) and the Brazilian Association of Zebu Breeders (Associação Brasileira de Criadores de Zebu). The database contained 50,402 TDMY records from 7113 first lactations of pure Guzerá cows, with records from 1987 to 2012 of calves in 101 herds that were born to 905 sires.

The milk yield records were divided into two test-day seasons, one corresponding to the rainy season (October to March) and the other to the dry season (April to September). The contemporary groups used in the analyses were categorized by herd, year, and season of test. Contemporary groups with fewer than three observations were eliminated. In addition, records with ± 3.5 SDs from the mean of the contemporary group were excluded. After applying these restrictions, the final number of contemporary groups was 918. After confirmation of the consistency of the database, there were 34,193 monthly production records of 5274 pure Guzerá cows born to 628 sires and 3302 dams. The relationship matrix contained 8589 animals.

The TDMY records were divided into 10 monthly classes, designated TDMY1 to TDMY10, which were obtained between days 6 and 305 of lactation. The numbers of observations in each class were 3205, 4205, 4161, 4008, 3881, 3620, 3431, 3050, 2546, and 2086 for TDMY1 to TDMY10, respectively. The lowest numbers of observations were obtained for the last two months of lactation.

Monthly milk yields were modeled using a single-trait RRM. The model included direct additive genetic, permanent environmental, and residual effects as random effects. Contemporary group was included as a fixed effect, and the linear and quadratic effects of age at calving and the number of days in milk (represented by the average lactation curve), respectively, were included as covariates. The average trajectory was modeled using a third-order Legendre polynomial, i.e., it contained four parameters.

The additive genetic and permanent environmental random effects were modeled using orthogonal Legendre polynomials with four and five parameters, respectively. The structure of residual variances was considered heterogeneous, and was divided into the following seven classes of variance: 30, 60, 90, 120-150, 180-210, 240-270, and 305 days in milk. This model was chosen after six different models had been tested using Akaike (AIC) and Bayesian Schwarz (BIC) (Wolfinger, 1993) criteria. The matrix representation of the model is:

$$y = Xb + Za + Wpe + e$$

where y is the vector of TDMY, b is the vector of solutions for fixed effects that included the solutions for contemporary group and for the covariates age at calving and days in milk, a is the vector of solutions for additive genetic random effects, pe is the vector of solutions for permanent environmental random effects, e is the vector of the random residual effect, X, Z, and W are the incidence matrices for fixed effects and the random effect of animal and permanent environment, respectively.

The model assumptions were as follows: $E_{(A)} = 0$, $E_{(PE)} = 0$ and $E_{(E)} = 0$, and $V_{(A)} = k_A \otimes A$, $V_{(PE)} = k_{PE} \otimes I_{ND}$ and $V_{(E)} = R$, where k_A and k_{PE} are (co)variance matrices between additive genetic and permanent environmental random regression coefficients, respectively; A is the relationship matrix between individuals; I_{ND} is the identity matrix of dimension Nd; $\otimes$ is the Kronecker product between matrices, and R is a diagonal block matrix containing residual variances. The variance components were estimated by the restricted maximum likelihood method using the WOMBAT statistical package (Meyer, 2006).

RESULTS AND DISCUSSION

An improvement in the Log L value was observed with an increasing number of parameters included in the model (Table 1). Model Leg 5.6.7 provided the best Log L and AIC values. The model with the best BIC value was Leg 5.5.7; however, model Leg 4.5.7 was chosen for the subsequent analysis, because this model had a BIC value that was similar to the best models, despite containing fewer parameters. Results from previous studies have demonstrated that although higher-order polynomials usually improve model plausibility, there are several problems associated with them. Additive genetic variances exhibit oscillatory patterns, which lead to extreme values at the peripheries of lactation and a negative correlation for the extremes of lactation (López-Romero et al., 2004; Bignardi et al., 2009a).

The overall mean TDMY was 7.36 ± 3.59 kg, with peak production occurring near the beginning of lactation (9.04 ± 4.62 kg). Production was almost the same in the second month of lactation as the first (8.98 ± 4.14 kg), and decreased in the third month until the end of lactation, when milk yield was 5.44 ± 2.53 kg. The coefficients of variation for TDMY were high throughout lactation, indicating differences in management, in lines within the breed, and in the genetic potential of each herd. The lactation peak (as early as the first month of lactation) agrees with previous results for Zebu breeds, in which the lactation curve is subdivided into only two segments:

early production, with a peak immediately at the beginning of lactation, or even no evident peak; and a decrease in milk yield until the end of lactation (Cobuci et al., 2000; Herrera et al., 2008).

Table 1. Number of parameters and statistical criteria used for the comparison of models employing Legendre polynomials.

Model	Statistical criteria			
	P	Log L	AIC	BIC
Leg 4.4.10	30	-37,165.1	37,195.09	37,321.27
Leg 4.4.7	27	-37,166.1	37,193.11	37,306.68
Leg 4.5.7	32	-37,108.1	37,140.12	37,274.71
Leg 4.6.7	38	-37,089.9	37,127.92	37,287.76
Leg 5.5.7	37	-37,070.9	37,107.89	37,263.52
Leg 5.6.7	43	-37,053.3	37096.34	37,277.21

Approximately 8% of the 5274 animals had one or two test-day records, whereas most animals (about 80%) had five or more. Initially, animals with fewer than two observations during lactation and those with yields well above the average for the breed were eliminated; however, since there was no significant difference in variance we chose not to exclude these records from the analysis.

The additive genetic variances in TDMY1 were high, with a sharp decrease on subsequent test days (Figure 1). The additive genetic variances stagnated by the fourth month, until the end of lactation. A similar trend was seen for the phenotypic variances. The highest permanent environmental variance was observed in TDMY1 and the lowest in TDMY9. This oscillation in permanent environmental variances may be explained by the fact that the cows in this dataset were all primiparous. Therefore, some cows had still not adapted to the milking management, and exhibited stress and difficulty in acclimatizing to the milking parlor.
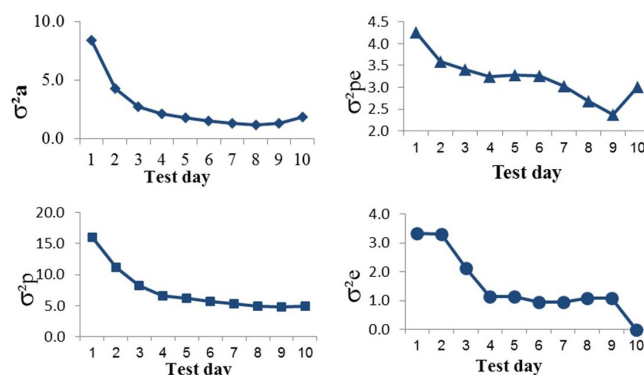
Figure 1. Additive genetic (σ^2_a), permanent environmental (σ^2_{pe}), phenotypic (σ^2_p), and residual (σ^2_e) variances between monthly test-day milk yields obtained with model Leg 4.5.7.

The residual variances were high during the first two months of lactation, and then decreased up to the fourth month before stabilizing. The residual variance was almost zero in the last month. Similar results were obtained by Brotherstone et al. (2000) and Bignardi et al. (2009a), who reported high residual variances at the beginning of lactation, probably because some cows have difficulty in adapting to the milking routine, in addition to post-partum stress and health problems.

The high additive variance in TDMY1 can be explained by the wide variation in the genetic potential of Guzerá animals, which are selected at different levels of intensity in breeding programs,

progeny tests, or multiple ovulation and embryo transfer nucleus breeding schemes. In addition, first-lactation cows with low production in the first month are often removed from official tests, and this may have occurred in the present study. The greater variation during this phase reflects the wide genetic variability in the shape of the curve. The additive and phenotypic variances exhibited similar trends at the end of lactation, with stagnation after the sixth month. This may be because only cows with long lactations and similar yields during this period were included in the dataset analyzed.

The heritabilities of TDMY (Table 2) ranged from 0.24 to 0.52, with the highest estimates in TDMY1. The heritabilities decreased gradually after the first month until TDMY8, and increased slightly on subsequent test days. Santos et al. (2013) used third- and sixth-order Legendre polynomials for genetic and permanent environmental effects, respectively, in Guzerá cattle, and estimated high heritabilities in TDMY2 (0.35) and low heritabilities in TDMY8 (0.18). The trends and magnitudes of the estimates are similar to those observed in the present study after the second month.

Table 2. Estimates of heritability and standard error (diagonal) and genetic (above the diagonal) and phenotypic (below the diagonal) correlations between test-day milk yields (TDMY).

Class	TDMY1	TDMY2	TDMY3	TDMY4	TDMY5	TDMY6	TDMY7	TDMY8	TDMY9	TDMY10
TDMY1	**0.52** (0.04)	0.95	0.80	0.61	0.47	0.39	0.34	0.29	0.18	-0.03
TDMY2	0.69	**0.38** (0.03)	0.94	0.82	0.70	0.62	0.54	0.44	0.29	0.05
TDMY3	0.60	0.69	**0.33** (0.03)	0.97	0.89	0.82	0.72	0.59	0.40	0.15
TDMY4	0.54	0.65	0.75	**0.32** (0.03)	0.98	0.93	0.84	0.70	0.50	0.25
TDMY5	0.46	0.56	0.68	0.80	**0.29** (0.03)	0.98	0.92	0.79	0.60	0.35
TDMY6	0.41	0.49	0.62	0.76	0.81	**0.26** (0.03)	0.97	0.88	0.72	0.47
TDMY7	0.37	0.44	0.56	0.69	0.76	0.81	**0.24** (0.03)	0.96	0.84	0.61
TDMY8	0.32	0.39	0.49	0.60	0.66	0.73	0.78	**0.24** (0.03)	0.95	0.77
TDMY9	0.26	0.33	0.42	0.50	0.56	0.62	0.69	0.74	**0.27** (0.04)	0.92
TDMY10	0.18	0.23	0.31	0.40	0.45	0.50	0.55	0.61	0.74	**0.38** (0.05)

The most desirable shape of the heritability curve is with low values at the beginning and end of lactation, as these periods are strongly influenced by non-genetic effects that accumulate before calving and are associated with decisions regarding drying-off (Druet et al., 2003). For Zebu cattle with short lactations, the number of records at the end of lactation may increase heritability in this period, and provide estimates of negative genetic correlations between production at the beginning and at the end of lactation, as observed by Costa et al. (2008).

High heritabilities were observed during early lactation in this study, because many animals were removed from the official dairy control at the beginning of lactation due to their low production. The heritability estimates exhibited a similar trend to that observed for the additive genetic variances, with high estimates at the beginning and end of lactation and lower values during mid-lactation. This finding agrees with Santos et al. (2013), who found high heritabilities on the test days when additive genetic variances were high. The fact that the heritability estimates obtained were of moderate to high magnitude throughout lactation, with high estimates at the extreme points, indicates that TDMY could contribute to selection for milk yield in Guzerá cattle, and that importance should be attached to yields at the beginning of lactation when using RRMs.

The genetic correlations between the TDMYs ranged from -0.3 to 0.98 (Table 2). High estimates were observed between adjacent test days, with strong correlations between TDMY1 and TDMY2 and between TDMY4 and TDMY5. In contrast, weak correlations were observed between distant TDMYs, and the only negative genetic correlation was between TDMY1 and TDMY10.

Strong genetic correlations always occur between adjacent TDMYs, and weak correlations are always observed as the interval between test days increases (Pereira et al., 2013; Santos et

al., 2013). The genetic correlations found in the present study are similar to those reported by Santos et al. (2013) for Guzerá cattle; those authors also found negative correlations between the first and last test-day months using RRMs.

The negative additive genetic correlation between the first (TDMY1) and the last (TDMY10) month may be related to the high heritabilities that were obtained for the first two test days, when only the first test-day records were available for a large number of cows, which resulted in negative correlations between early and more distant test days. Although very weak, the negative genetic correlation between TDMY1 and TDMY10 suggests that selection for high yields at the beginning of lactation leads to antagonistic selection for milk yield at the end of lactation, and, consequently, for lower lactation persistency. This conclusion is supported by the medium to weak correlations observed between the TDMYs obtained during the second part of lactation and the early yields (TDMY1 and TDMY2).

The genetic correlations suggest that most of the genetic variation in milk yield can be explained by genes that act throughout the period of lactation, indicating that selection for a particular stage of lactation will increase yield in the other stages, except for TDMY1 and TDMY2 with TDMY10.

The phenotypic correlations were all positive, and ranged from 0.18 (TDMY1 and TDMY10) to 0.81 (TDMY6 and TDMY7); the values decreased with increasing intervals between test days (Table 2). The permanent environmental correlations were all positive, and ranged from 0.48 to 0.98, with the lowest estimates being between distant TDMYs (data not shown). The lowest permanent environmental correlation was observed between TDMY2 and TDMY10 (0.48), whereas the highest estimates were obtained between adjacent test days, such as TDMY7 and TDMY6 (0.98). The permanent environmental correlations gradually weakened with increasing distance between test days, as expected. In general, the permanent environmental correlations were always stronger than the genetic correlations, except between TDMY1 and TDMY2.

The results of the present study confirm the suitability of RRMs as a complement to traditional models for 305-day cumulative milk yield, because selection based on TDMYs obtained with RRMs also selects for the shape of the individual lactation curve. Therefore, the use of RRMs is an effective option for the evaluation of Zebu animals that exhibit low lactation persistency and short lactations. Furthermore, RRMs can be used to estimate breeding values for strategic points or intervals of lactation, and for lactation persistency. However, the number of official milk yield records for Zebu animals in Brazil is still low. Therefore, genetic parameter estimates for Guzerá cattle may change significantly as new records become available.

CONCLUSIONS

The genetic parameters for TDMYs by RRMs agree with those reported in the literature, and indicate reasonable genetic variability in milk yield during the different phases of lactation, with the greatest variability at the beginning of lactation. The general trend of the genetic correlations for milk yield on different test days indicates that selection for milk yield at a particular stage of lactation will increase production during the other stages, but at a low intensity during the final stages.

ACKNOWLEDGMENTS

REFERENCES

Bignardi AB, El Faro L, Albuquerque LG, Cardoso VL, et al. (2009a). Random regression models to estimate test-day milk yield genetic parameters Holstein cows in southeastern Brazil. *Livest. Sci.* 123: 1-7.

Bignardi AB, El Faro L, Cardoso VL, Machado PF, et al. (2009b). Parametric correlation functions to model the structure of permanent environmental (co)variances in milk yield random regression models. *J. Dairy Sci.* 92: 4634-4640.

Brotherstone S, White IMS and Meyer K (2000). Genetic modeling of daily yields using orthogonal polynomials and parametric curves. *J. Anim. Sci.* 70: 407-415.

Cobuci JA, Euclydes RF, Verneque RS, Teodoro RL, et al. (2000). Curva de lactação na raça Guzerá. *Rev. Bras. Zootec.* 29: 1332-1339.

Costa CN, Melo CNR, Pacher IU, Freitas AF, et al. (2008). Genetic parameters for test-day milk yield of first lactation Holstein cows estimated by random regression using Legendre polynomials. *Rev. Bras. Zootec.* 37: 602-608.

Druet T, Jaffrezic F, Boichard D and Ducrocq V (2003). Modeling lactation curves and estimation of genetic parameters for first lactation test-day records of French Holstein cows. *J. Dairy Sci.* 86: 2480-2490.

Herrera LGG, El Faro L, Albuquerque LG, Tonhati H, et al. (2008). Estimativas de parâmetros genéticos para a 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.

López-Romero P, Rekaya R and Carabano MJ (2004). Bayesian comparison of test-day models under different assumptions of heterogeneity for the residual variance: the change point technique versus arbitrary intervals. *J. Anim. Breed. Genet.* 121: 14-25.

Meyer K (1998). Estimating covariance functions for longitudinal data using a random regression model. *Genet. Sel. Evol*. 30: 221-240.

Meyer K (2006). "WOMBAT"- digging deep for quantitative genetic analyses by restricted maximum likelihood. Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brazil (CD-ROM).

Peixoto MGCD, Poggian CF, Verneque RS, Egito AA, et al. (2010). Genetic basis and inbreeding in the Brazilian Guzerat (*Bos indicus*) subpopulation selected for milk production. *Livest. Sci.* 131: 168-174.

Pereira RJ, Bignardi AB, El Faro L, Verneque RS, et al. (2013). Random regression models using Legendre polynomials or linear splines for test-day milk yield of dairy Gyr (*Bos indicus*) cattle. *J. Dairy Sci.* 96: 565-574.

Santos DJ, Peixoto MG, Borquis RR, Verneque RS, et al. (2013). Comparison of random regression models to estimate genetic parameters for milk production in Guzerat (*Bos indicus*) cows. *Genet. Mol. Res.* 12: 143-153.

Wolfinger R (1993). Covariance structure selection in general mixed models. *Commun. Stat. - Simul. C* 22: 1079-1106.