

Estimation of genetic parameters for milk yield in Murrah buffaloes by Bayesian inference

F. C. Breda,*¹ L. G. Albuquerque,†‡ R. F. Euclides,§ A. B. Bignardi,† F. Baldi,† R. A. Torres,§ L. Barbosa,# and H. Tonhati†‡²

*Universidade Federal de Roraima (UFRR), 69300-000, Boa Vista, RR, Brazil

†Universidade Estadual Paulista (FCAV/UNESP), 14884-900, Jaboticabal, SP, Brazil

‡Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Instituto Nacional de Ciência e Tecnologia - Ciência Animal (INCT-CA[**AU1: Add city, postal code and country**])

§Universidade Federal de Viçosa (UFV), 36571-000, Viçosa, MG, Brazil

#Universidade de Sergipe (UFS), 49037-470, Aracaju, SE, Brazil

ABSTRACT

Random regression models were used to estimate genetic parameters for test-day milk yield in Murrah buffaloes using Bayesian inference. Data comprised 17,935 test-day milk records from 1,433 buffaloes. Twelve models were tested using different combinations of third-, fourth-, fifth-, sixth-, and seventh-order orthogonal polynomials of weeks of lactation for additive genetic and permanent environmental effects. All models included the fixed effects of contemporary group, number of daily milkings and age of cow at calving as covariate (linear and quadratic effect). In addition, residual variances were considered to be heterogeneous with 6 classes of variance. Models were selected based on the residual mean square error, weighted average of residual variance estimates, and estimates of variance components, heritabilities, correlations, eigenvalues, and eigenfunctions. Results indicated that changes in the order of fit for additive genetic and permanent environmental random effects influenced the estimation of genetic parameters. Heritability estimates ranged from 0.19 to 0.31. Genetic correlation estimates were close to unity between adjacent test-day records, but decreased gradually as the interval between test-days increased. Results from mean squared error and weighted averages of residual variance estimates suggested that a model considering sixth- and seventh-order Legendre polynomials for additive and permanent environmental effects, respectively, and 6 classes for residual variances, provided the best fit. Nevertheless, this model presented the largest degree of complexity. A more parsimonious model, with fourth- and sixth-order polynomials,

respectively, for these same effects, yielded very similar genetic parameter estimates. Therefore, this last model is recommended for routine applications.

Key words: covariance function, genetic correlation, heritability, test-day milk yield

INTRODUCTION

The world buffalo milk production increased 43% from 1997 to 2007. This rate is superior to that reported for cow milk production (20%) in the same period (FAO, 2009). However, buffalo milk yield is still much lower than cow's milk yield. In 2007, worldwide buffalo and cow milk yields were 85 and 560 million tons[**AU2: Metric tonnes?**], respectively (FAO, 2009). Higher average milk yields are reported in India and Italy, probably because genetic evaluations are a common practice in these countries (Moioli and Borghese, 2005).

Few studies have reported milk yield genetic parameter estimates in buffaloes. These are generally restricted to total milk yield, with heritability estimates ranging from 0.14 to 0.40 (Rosati and Van Vleck, 1998; Tonhati et al., 2000; Peeva, 2002), and test-day milk yield, with estimates ranging from 0.01 to 0.24 (Hurtado-Lugo et al., 2006; Aspilcueta-Borquis et al., 2007).

Random regression models have become a standard procedure for the genetic analysis of longitudinal traits such as milk yield (Meyer, 2005) because of their flexibility and ability to describe (co)variances among test-day milk yield records measured on different days during the lactation (Jensen, 2001). Random regression models allow estimates of covariances between coefficients of random functions or, equivalently, estimates of covariance functions.

Compared with REML, Bayesian methods have the advantage of allowing the inclusion in analysis of prior knowledge about the unknown parameters. Additionally, Bayes theorem provides a solution for the finite sample size problem because an exact distribution of

Received March 20, 2009.

Accepted October 7, 2009.

¹F. C. Breda acknowledges Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Universidade Federal de Roraima (UFRR) for financial support.

²Corresponding author: tonhati@fcav.unesp.br

Table 1. Summary of data structure after edits

Records, n	17,935
Animals with records, n	1,433
Sires, n	114
Dams, n	975
Contemporary groups, n	708
Average daily milk yield, kg	6.25
Standard deviation of daily milk yield, kg	2.29

the posterior exists for each large or small data set from which inferences can be drawn. When a large data set is analyzed, a priori information tends to be overwhelmed by the likelihood function in the establishment of the a posteriori distribution. In this case, parameter estimates are close to those obtained by frequentist methods based on likelihood functions. However, this may not be true when the sample size is limited because the maximum likelihood procedure only possesses well-defined properties when the sample size is large enough (Gianola and Fernando, 1986).

The objective of the present study was to estimate covariance functions for additive genetic and permanent environmental effects and genetic parameters for test-day milk yield in Murrah buffaloes using random regression models and Bayesian inference.

MATERIALS AND METHODS

Experimental Setup and Analyses

A total of 17,935 test-day milk yield records from 1,433 first lactations of Murrah buffaloes, raised in 11 herds in São Paulo state in Brazil and calving from 1985 to 2006, were analyzed. Cows' age ranged from 24 to 48 mo. Milk records were obtained from d 7 of lactation onward and lactation was truncated at 305 d. Test-days were grouped into weekly classes of days in milk, ranging from 2 to 43 wk of lactation. The first test-day milk record was required to be at the latest 75 d after calving. In addition, records outside the range of 3 standard deviations from the mean were excluded.

Contemporary groups (CG) were formed by farm-year-recording month subclasses. Only data of animals with at least 4 test-day records during lactation and CG with at least 5 animal observations were kept. A total of 708 CG were formed. The average number of animals per CG was 27 (ranging from 5 to 66). The relationship matrix had a total of 2,810 animals. The structure of the data set after editing is summarized in Table 1.

Preliminary analyses using least squares were performed to determine the effects of age of cow at calving, number of daily milkings and week of lactation on the trait studied. The additive genetic and permanent environmental effects were considered as random effects in

the model and were modeled by Legendre polynomials (LP) of weeks of lactation considering different orders.

Estimates of the (co)variance components were obtained by Bayesian analysis using the RRGIBBS program (Meyer, 2002). This program generates Markov chains for the parameters of a random regression model by Gibbs sampling. For each analysis 5 chains with different starting values were run. After several trials, the length of the chain was set to 400,000. The burn-in period was 100,000 iterations, higher than the minimum burn-in required according to the method of Raftery and Lewis (1992). Convergence was tested using the criteria proposed by Heidelberger and Welch (1983) and Geweke (1992) criteria. The software R, with some routines of the package Bayesian Output Analysis (BOA), was used to calculate Geweke's and Heidelberger and Welch's statistics (Smith, 1997).

The following variables were included in all random regression models analyzed: contemporary group (708 levels) and milking number (2 levels) as fixed effects, linear and quadratic effects of cow age at calving as covariate (linear and quadratic effect), and the average trajectory of the population modeled by Legendre polynomials of third order. The matrix presentation of the single-trait random regression model is given by

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wc} + \mathbf{e},$$

where $\mathbf{y}$ is the vector of N observations measured in N_d animals; $\mathbf{b}$ is the vector of the systematic effects and fixed regression coefficients; $\mathbf{a}$ is the vector of additive genetic random regression coefficients; $\mathbf{c}$ is the vector of permanent environmental random regression coefficients; $\mathbf{e}$ is the random vector of error effects; and $\mathbf{X}$, $\mathbf{Z}$, and $\mathbf{W}$ are the incidence matrices of fixed effects and additive genetic and permanent environmental random effects.

The following assumptions are made:

- $y | b, a, c, \sigma_{el}^2, \dots, \sigma_{ef}^2 \sim NMV(\mathbf{Xb} + \mathbf{Za} + \mathbf{Wc}, R)$, with $R = \text{diag}\{\sigma_{el}^2\}$, the likelihood function, in that $l = 1, 2, \dots, f$, where f is the number of residual classes.
- $b \propto \text{constant}$;
- $a | \mathbf{K}_a \sim NMV(0, G)$, with $G = \mathbf{A} \otimes \mathbf{K}_a$, where $\mathbf{A}$ is the numerator relationship matrix between animals and $\mathbf{K}_a$ is the matrix containing (co)variances between additive genetic random regression coefficients;
- $c | \mathbf{K}_c \sim NMV(0, C)$, with $C = \mathbf{I} \otimes \mathbf{K}_c$, where $\mathbf{I}$ is an identity matrix, $\mathbf{K}_c$ is the matrix containing (co)variances between permanent environmental

random regression coefficients, and *NMV* refers to a multivariate normal distribution;

- $\mathbf{K}_a | v_a, S_a^2 \sim W^{-1}(v_a, v_a S_a^2)$;
- $\mathbf{K}_c | v_c, S_c^2 \sim W^{-1}(v_c, v_c S_c^2)$, where v_a, S_a^2 and v_c, S_c^2 are interpreted as degrees of belief and priori values for (co)variances of the additive genetic and permanent environmental regression coefficients, respectively, and W^{-1} is an inverted Wishart distribution.
- $\sigma_{el}^2 | v_{el}, s_{el}^2 \sim X^{-2}(v_{el}, v_{el} s_{el}^2)$, where X^{-2} is a scaled inverse chi-square distribution.

Twelve models were tested using different combinations of third-, fourth-, fifth-, sixth-, and seventh-order Legendre polynomials (**LP**) of weeks of lactation for additive genetic and permanent environmental effects. Therefore, models were identified as follows: LP(*ij*) where *i* and *j* indicate the order of fit for additive genetic and permanent environmental effects, respectively.

Preliminary analyses were conducted to determine the most adequate structure to model the residual variance. Step functions were initially fitted with 42 weekly classes of residual variances. Afterward, the pattern of variation was analyzed to define step functions with 4, 6, or 22 residual variance classes based on residual variances estimated using the initial 42 classes. In short, weeks of lactation were grouped in the following classes: 4 (2, 3–4, 5, 6–43 wk), 6 (2, 3–4, 5, 6–14, 15–34, 35–42 wk), 22 (2, 3, 4, 5, 6, 7, 8, 9, 10, 11–14, 15–19, 20–25, 26–29, 30, 31, 32–33, 34–36, 37–38, 39, 40–41, 42–43 wk), or 42 weekly classes. A model with homogeneous residual variance across lactation was also tried. These preliminary analyses (results not shown) showed that the best structure to model the residual variance was a step function with 6 classes. This structure was used in all remaining analyses.

The highest posterior density interval was determined for each parameter of the model with 95% confidence. The residual mean square (**MSR**), weighted average of residual variance estimates (**RES**), and estimates of genetic parameters were used as criteria to choose the most adequate model to describe changes in additive genetic, permanent environmental, and residual variances along the lactation curve. In addition, the degree of interdependence between model parameters (i.e., complexity) was also calculated and is given by the second term of the information-theoretic measure of complexity (Bozdogan, 2000). The RES and measure of complexity were calculated as follows:

$$RES = \frac{1}{TS} \sum_{l=1}^f \hat{\sigma}_e^2 NS_l,$$

where $\hat{\sigma}_{el}^2$ = residual variance; *f* = number of residual classes; NS_l = number of weeks in the classes; and TS = total number of weeks (Jamrozik and Schaeffer, 2002).

$$C_{Modelo} = C_{\mathbf{K}_a} + C_{\mathbf{K}_c},$$

with

$$C_{\mathbf{K}_a} \text{ or } C_{\mathbf{K}_c} = \frac{p(\mathbf{K})}{2} \ln \left[\frac{\text{trace}(\mathbf{K})}{p(\mathbf{K})} \right] - \frac{1}{2} \ln |\mathbf{K}|,$$

where $p(\mathbf{K})$ is the rank of matrix $\mathbf{K}$. Greater simplicity is reached when $\mathbf{K}_a$ and $\mathbf{K}_c$ tend toward an identity matrix, suggesting that the parameters are orthogonal and can be estimated with equal precision.

Standard Single-Trait Model

To assess the pattern of changes in variances along the lactation and to compare with results obtained with random regression models, the same data set was analyzed with a series of single-trait models. Thus, 9 separate single-trait standard analyses were performed, considering milk yield from d 7 to 60, 30 to 90, 60 to 120, 90 to 150, 120 to 180, 150 to 210, 180 to 240, 210 to 270, and 240 to 300, respectively. The model included the fixed effects of contemporary group (herd-year-recording month), number of milkings, and age of cow at calving as covariate (linear and quadratic effects). The additive genetic effect was modeled by fitting a random animal effect.

RESULTS AND DISCUSSION

The number of records at each test week and the corresponding mean of milk yield are shown in Figure 1. The peak period of milk production of buffaloes was observed around the 11th week, when animals produced an average of 7.30 kg of milk per test day. After this period, milk yield decreased with increasing days in milk. The number of records was smaller at either extreme of the lactation curve (Figure 1). The standard deviation of milk yield ranged from 2.25 kg (wk 2) to 1.45 kg (wk 43).

The MSR, percentiles of RES, and measures of complexity for each model converged according to Heidelberg and Welch (1983) and Geweke (1992) criteria are shown in Table 2. The degree of complexity tended to increase with the number of parameters to

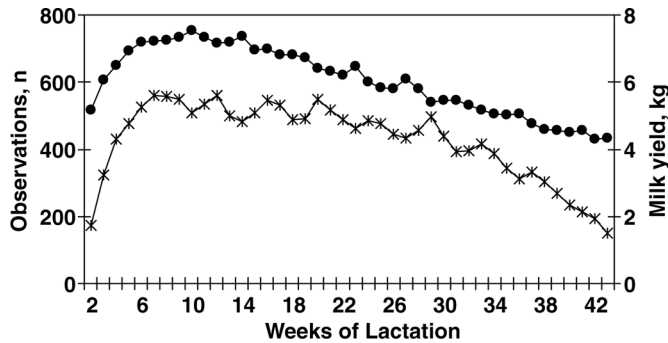


Figure 1. Number of observations (*) and average daily milk yield (●) according to test week of first lactations of Murrah buffaloes.

be estimated. Model LP67, which included the largest number of parameters to be estimated, had the lowest MSR and RES. However, the number of parameters was not the main factor that influenced the reduction in the residual variance. For example, low and similar estimates of MSR and RES were obtained for models LP36, LP46, LP56, LP65, and LP67, that included from 33 to 55 parameters.

Irrespective of the order of fit for permanent environmental effect, genetic variance estimates remained almost unchanged (data not shown) when a third-order polynomial was fitted for the genetic effect. According to other studies (López-Romero et al., 2003; Roos et al., 2004; Assis et al., 2006), this result is biologically unlikely. In addition, these models were associated with high MSR and RES estimates. Models fitting polynomials of order >3 (from 4 to 6) for the genetic effect led to more acceptable genetic variance estimates.

Genetic variance estimates obtained with LP65 model were higher at the beginning of lactation when compared with a model containing the same number

of parameters (LP56). Also, heritability estimates obtained with LP65 model showed marked fluctuations along the lactation curve (data not shown). This finding suggests that the use of an order of fit for permanent environmental effect smaller than those applied for the additive genetic effect may not be adequate to model changes in (co)variances along the lactation trajectory. Similar results were also reported for dairy cattle and goats by López-Romero and Carabaño (2003), Assis et al. (2006), and Bignardi et al. (2009).

Variance Components and Heritability

The posterior distribution features of the (co)variances estimates between random regression coefficients for all models were similar. The deviation of the posterior mean in relation to respective median (50% percentile) provides information regarding the symmetry of the marginal posterior distributions. In general, mean and median estimates for the additive genetic and permanent environmental random coefficients and the residual variance classes were similar.

Serial correlation provides information regarding the degree of association between samples in the Markov chain. In the present study, serial correlations were high, with the highest correlations being observed for the additive genetic random regression coefficients. This resulted in a smaller effective size of the sample for this effect when compared with permanent environmental effects. Similar results have been reported by López-Romero et al. (2003). Another aspect is that the models including a large number of parameters presented difficulty to converge to the stationary distribution, requiring a large number of samples and a longer burn-in period the first time the Markov chain was generated.

Table 2. Fitted models and their respective number of estimated parameters (p), residual mean squared (MSR), percentiles of weighted average of residual variance estimates (RES), and measure of complexity

Model ¹	p	MSR	RES			Complexity
			2.5%	50%	97.5%	
LP33	18	0.91	1.01	1.15	1.22	3.25
LP34	22	0.84	0.96	1.07	1.13	4.25
LP35	27	0.78	0.90	1.01	1.06	6.07
LP36	33	0.74	0.86	0.96	1.00	7.79
LP44	26	0.83	0.94	1.05	1.11	6.58
LP45	31	0.78	0.89	1.00	1.05	8.14
LP46	37	0.73	0.85	0.95	1.00	9.72
LP54	31	0.78	0.88	0.98	1.03	7.91
LP55	36	0.77	0.88	0.98	1.02	11.38
LP56	42	0.73	0.84	0.94	0.98	12.79
LP65	42	0.74	0.84	0.94	0.98	12.80
LP67	55	0.71	0.81	0.91	0.96	16.99

¹LP(*ij*), where LP = Legendre polynomials; *i* and *j* indicate the order of fit for additive genetic and permanent environmental effects, respectively.

Models LP46, LP56, and LP67 differed little in the partition of phenotypic variance into genetic, permanent environmental, and residual variance. The largest differences were observed at the beginning and end of lactation. Estimates obtained with models LP56 and LP67 were very similar to each other. Phenotypic variances obtained by the 3 models were very similar to those obtained with standard single-trait analyses, ranging from 2.44 to 5.47 kg², and were higher at wk 2 followed by a strong decrease and remaining almost unchanged after wk 6. A similar behavior was observed for permanent environmental variances which were higher than the genetic variance estimates, except at the end of the lactation curve. Genetic variances obtained with random regression models showed the same trend as those estimated by a single-trait model (data not shown).

The magnitude of variance not explained by the model (σ_e^2) decreased with week of lactation, with residual variance estimates of 1.30, 1.23, 1.02, 0.88, and 0.71 kg². A similar trend was reported by Assis et al. (2006) for milking goats.

Whereas residual variance estimates decreased at the end of lactation, estimates of genetic variances increased. Cows with high persistency of lactation and similar milk yields at the end of lactation tend to remain in the herd. These animals are probably influenced in a similar manner by the environment (e.g., feeding, climate, management) and most of the differences in milk yield observed can be attributed to genetic rather than environmental factors.

Figure 2 shows the posterior means of heritability (h^2) and proportion of phenotypic variance corresponding to permanent environmental variance (c^2) estimated using the models LP46, LP56, and LP67. Changes in heritability estimates with week of lactation reflect the pattern described previously for variance component estimates and followed trends in genetic variance estimates. Differences in heritability estimates between models were detected at the beginning (0.33 and 0.44 for LP56 and LP67, respectively) and the end of lactation (0.43 and 0.50 for LP56 and LP46, respectively), when heritabilities were high, except for those obtained with model LP46 at the beginning of lactation (0.19). Furthermore, a small difference in heritability estimates obtained with the 3 models was observed during the phase of highest milk production (wk 6 to 14), with model LP56 yielding higher estimates than LP46 and lower estimates than LP67. After the second week, heritability declined until wk 6 (0.20), a period characterized by the onset of milk production peak (Figure 1). Heritability estimates increased between wk 6 and 12 and then decreased again until wk 31 (0.21). Ignoring

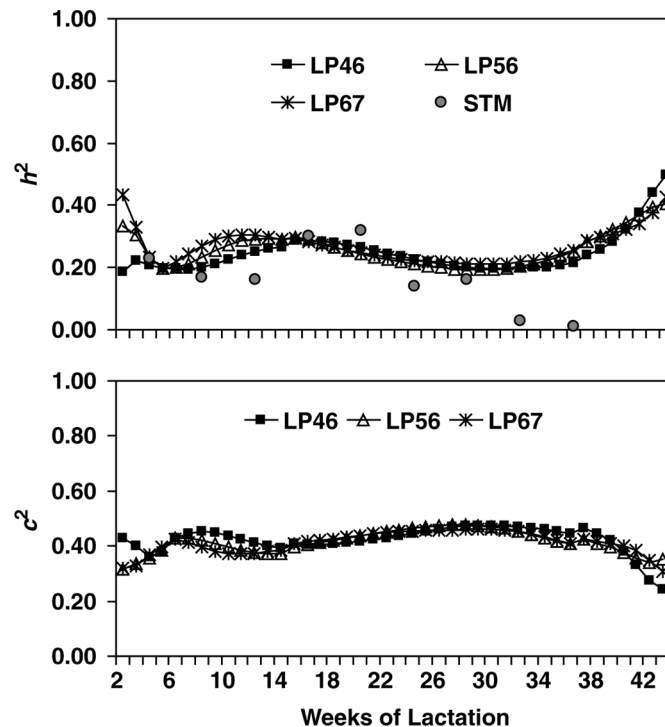


Figure 2. Posterior means of heritability (h^2) and the proportion of phenotypic variance corresponding to permanent environmental variance (c^2) for weekly test-day milk obtained with the LP46, LP56, and LP67 random regression models and by a standard single trait model (STM). LP(ij), where LP = Legendre polynomials and i and j indicate the order of fit for additive genetic and permanent environmental effects, respectively.

heritabilities for milk yield in the first and final weeks of lactation, the estimates ranged from 0.19 to 0.31. The highest estimates obtained by the 3 random regression models were between wk 9 and wk 19. Until the 28th week of lactation, the heritability estimates obtained with standard single-trait test-day models showed the same trend as the random regression model estimates.

As can be seen in Figure 2, the largest influence of the permanent environment effect was observed on wk 6 of lactation (0.43) and between wk 21 and 37 (0.45 to 0.48). These weeks were also characterized by a smaller genetic influence, indicating a strong sample correlation between h^2 and c^2 . Meyer (2005) observed a lower permanent environmental variance for models with a higher order of k_c . This trend was not observed in the present study when k_c was increased from sixth (LP46 and LP56) to seventh order (LP67). This finding suggests that an order of k_c equal to 6 is sufficient to model the changes in permanent environmental variances across lactation without producing marked alterations in the genetic and residual variance estimates.

The largest differences between random regression models in terms of variance component estimates, heri-

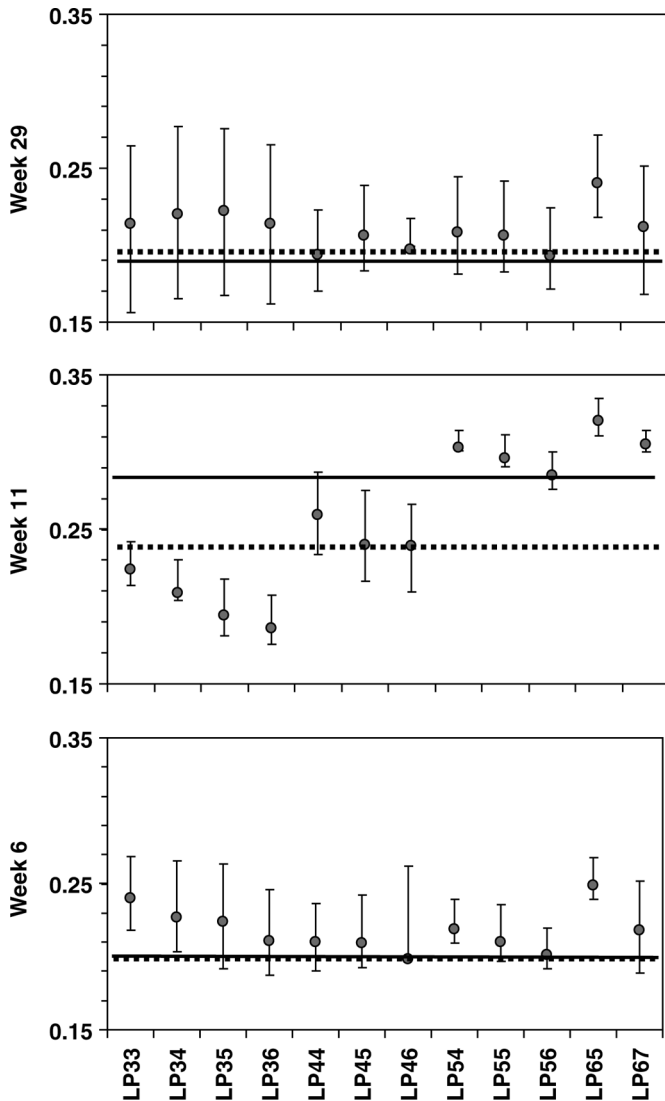


Figure 3. Posterior means of heritabilities (●) and approximate 95% highest posterior density regions (bars) of the models with different orders of fit of polynomials for wk 6, 11, and 29 of lactation: solid and dashed lines indicate the posterior mean of heritabilities estimated using models LP56 and LP46, respectively. LP(ij), where LP = Legendre polynomials and i and j indicate the order of fit for additive genetic and permanent environmental effects, respectively.

heritabilities, and c^2 were observed at the beginning and the end of the lactation. This finding can be explained by the small number of records available for analysis during these periods (Figure 1). The same finding was described by Meyer (2005) in a random regression analysis of growth traits in Angus. During the lactation peak (from wk 6 to 14), a period characterized by a large number of records, the models also showed differences in the estimates of σ_c^2 and σ_a^2 , which affected the heritability estimates (Figure 2). However, these differences were smaller than those observed at the beginning and end of lactation.

The difference between models is better illustrated in Figure 3, which shows the heritability estimates and their approximate confidence intervals for wk 6, 11, and 29. In general, for wk 6 and 29, models LP46 and LP56 presented posterior means of heritabilities that were close to and within the confidence interval observed in other models. In wk 11, a smaller confidence interval was observed for model LP56 than those obtained for LP46 model, suggesting higher reliability of the estimates provided by model LP56 at wk 11.

Correlations

Negative genetic correlations between the initial and final weeks of lactation were obtained for the 3 random regression models. However, models LP56 and LP67 yielded negative genetic correlations also between wk 2 and mid-lactation (after wk 9). Negative genetic correlations obtained with models LP56 and LP67 involving the second week of lactation might be attributed to the inability of high-order polynomials to model the extremes of the lactation curve. In addition, during this stage, the cow usually suffers from postcalving stress and energy deficits as reported by Bignardi et al. (2009) for dairy cattle. A similar trend has been also described in the literature (Brotherstone et al., 2000; López-Romero and Carabaño, 2003).

In general, except for the first week, estimates of genetic, permanent environmental, and phenotypic correlations decreased with increasing interval between test weeks. This behavior is illustrated in Figure 4, which shows the genetic and permanent environmental correlations between milk yield at wk 11 (peak production) and the other weeks. Genetic correlations obtained by model LP46 were high until wk 25 and then decreased rapidly over subsequent weeks. In contrast, estimates obtained with models LP56 and LP67 after wk 15 began to decrease slowly. The genetic correlation between wk 11 and 37 was 0.56 for LP56, 0.48 for LP67, and 0.43 for LP46, suggesting that selection for higher milk yields at the beginning of lactation may exert positive effects on milk yield during the final weeks (Olori et al., 1999; Assis et al., 2006). With respect to permanent environmental correlations, differences were observed during the first 4 wk, with lower estimates obtained with model LP46. Similar permanent environmental correlations were obtained with the 3 models in the other weeks.

Eigenvalues and Eigenfunctions

The estimates of the first 2 eigenvalues of the genetic and permanent environmental covariance functions obtained with LP46, LP56, and LP67 models were

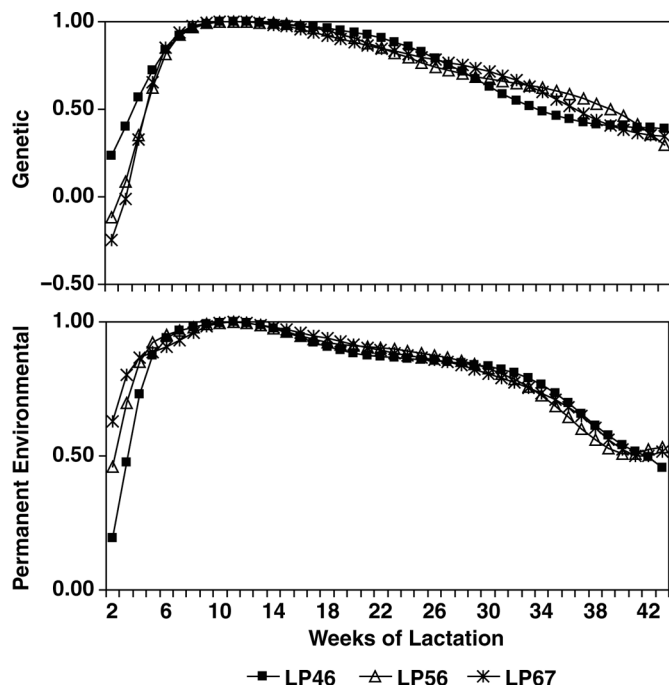


Figure 4. Posterior means of genetic and permanent environmental correlations between milk yield in wk 11 and wk 2 to 43. LP(ij), where LP = Legendre polynomials and i and j indicate the order of fit for additive genetic and permanent environmental effects, respectively.

practically the same: the first 2 eigenvalues of the coefficient matrix of the additive genetic covariance function corresponded to 88, 87, and 86% of total genetic variation, respectively, and the first 2 eigenvalues of the coefficient matrix of the permanent environmental covariance function accounted for approximately 89% of variation in each case.

The 2 main eigenfunctions for additive genetic effects obtained using models LP46, LP56, and LP67 are shown in Figure 5. The first eigenfunction obtained with model LP46 was positive and tended to oscillate along the trajectory of lactation. The first eigenfunctions obtained with LP56 and LP67 models were only negative in the second week. This finding probably explains the negative genetic correlations observed between the second week and mid lactation. In general, the first eigenfunction increased over the first weeks of lactation and only decreased after the lactation peak, remaining practically constant in the subsequent weeks. The relatively constant values observed after the period of production peak indicate that milk yield is mainly controlled by genes with similar effects (Olori et al., 1999). The positive first eigenfunction suggests that milk yield can change in the same direction in all weeks. In addition, this change is easily achieved because approximately 70% of the genetic variation is attributable to the first eigenfunction.

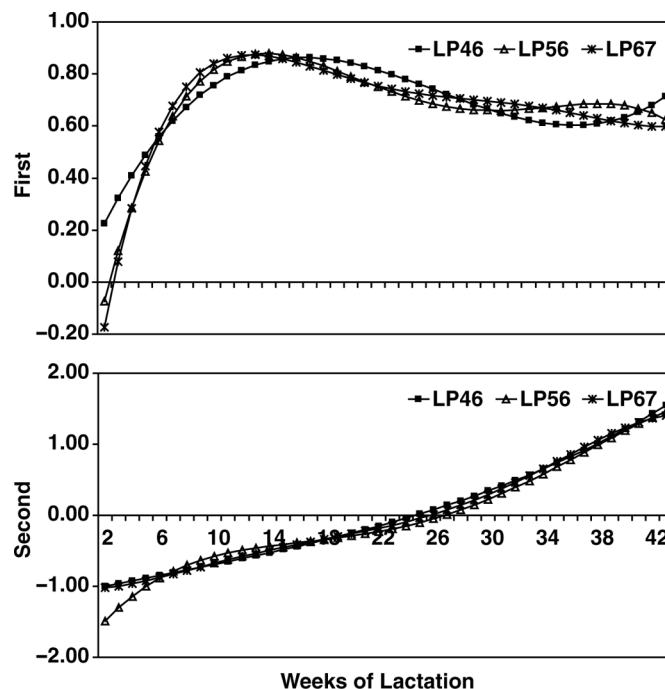


Figure 5. Estimates of the first and second genetic eigenfunctions obtained using models LP46, LP56, and LP67. LP(ij), where LP = Legendre polynomials and i and j indicate the order of fit for additive genetic and permanent environmental effects, respectively.

The second eigenvalue of the genetic covariance functions provided by models LP46, LP56, and LP67 corresponded to 18, 18, and 17% of the sum of eigenvalues, respectively. For all 3 models, the corresponding eigenfunction was negative until mid-lactation (wk 26) and positive in the following weeks. This sign change suggests an association with a factor that exerts contrasting effects on milk yield at the beginning and the end of lactation. Selection based on such a factor (e.g., persistency of lactation) might be used to change the shape of the lactation curve (Olori et al., 1999). The other eigenvalues together were responsible for less than 15% of the total sum, suggesting a small potential for change in milk yield pattern along lactation when selection is based on the third, fourth, fifth, and sixth eigenfunctions, as they contribute little to total genetic variation.

Results obtained in this study clearly illustrate the effects of the random regression model chosen on the estimates of genetic parameters. According to Miszta et al. (2000), there are marked differences between estimates of genetic parameters obtained with different random regression models. These differences might be attributed to the small number of milk yield records or to the models and functions used to describe the random regression. Meyer (2005) suggested that the choice of the best model is a compromise between the

complexity of the model, credibility of the results and fit (measured by the mean square error). In this respect, models LP46, LP56, and LP67 produced similar estimates of the genetic parameters, with the largest differences being observed at the beginning and end of lactation and small differences being found in the weeks corresponding to peak of milk production. An order of $k_c = 7$ did not change the variance estimates when compared with models that used $k_c = 6$, thus not justifying the adoption of this model to describe changes in (co)variances along the lactation curve. In addition, models LP67 and LP56 resulted in a higher degree of interdependence between additive genetic and permanent environmental random regression coefficients, with a consequent higher estimated complexity of the model as illustrated in Table 2.

The LP46 is a relatively parsimonious model; it resulted in a smaller degree of interdependence between random regression coefficients and yielded genetic parameter estimates similar to LP67 and LP56 models.

CONCLUSIONS

The fitted random regression models detected genetic variability and the existence of positive genetic correlations between test-day milk yields, which suggests that the selection of buffaloes that are genetically superior for milk yield on a given test-day may result in positive responses at any point of the lactation curve. A model with orders of 4 and 6 for genetic additive and permanent environmental effects is recommended to describe changes in milk yield (co)variances along the lactation curve for this data set.

REFERENCES

- Aspilcueta-Borquis, R., A. Tanaka, L. G. Albuquerque, et al. 2007. Genetic parameters for milk, fat and protein yield analyzed by test day models for Murrah buffaloes. *Proc. 8th World Buffalo Congress*, Caserta, Itália. **[AU3: Please provide names of all authors; provide name and location of proceedings publisher]**
- Assis, G. M. L., L. G. Albuquerque, J. L. R. Sarmento, J. M. Carneiro Junior, P. S. Lopes, and M. T. Rodrigues. 2006. Estimación de parámetros genéticos em caprinos leiteiros por meio de análise de regressão aleatória utilizando-se a amostragem de Gibbs. *R. Bras. Zootec.* 35:706–714.
- Bignardi, A. B., L. El Faro, V. L. Cardoso, P. F. Machado, and L. G. Albuquerque. 2009. Random regression models to estimate test-day milk yield genetic parameters Holstein cows in Southeastern Brazil. *Livest. Sci.* 123:1–7.
- Bozdogan, H. 2000. Akaike's information criterion and recent developments in information complexity. *J. Math. Psychol.* 44:62–91.
- Brotherstone, S., I. M. S. White, and K. Meyer. 2000. Genetic modeling of daily milk yield using orthogonal polynomials and parametric curves. *J. Anim. Sci.* 70:407–415.
- Food and Agriculture Organization (FAO). 2009. <http://faostat.fao.org>. Accessed June 01, 2009.
- Geweke, J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. Pages 625–631 in *Bayesian Statistics 4* (eds. Bernardo, J.M.; Berger, J.O.; Dawid, A.P.; Smith, A.F.M.), New York, Oxford University Press.
- Gianola, D., and R. L. Fernando. 1986. Bayesian methods in animal breeding theory. *J. Anim. Sci.* 63:217–244.
- Heidelberger, P., and P. Welch. 1983. Simulation run length control in the presence of an initial transient. *Oper. Res.* 31:1109–1144.
- Hurtado-Lugo, N., M. Cerón-Muñoz, and A. Gutiérrez-Valencia. 2006. Estimación de parámetros genéticos para la producción de leche en el día del control en búfalos de la Costa Atlántica Colombiana. *Livest. Res. Rural Devel.* 18:3. **[AU4: Please provide the last page number if article is longer than one page.]**
- Jamrozik, J., and L. R. Schaeffer. 2002. Bayesian comparison of random regression models for test-day yields in dairy cattle. *Proc. 7th World Congress of Genetics Applied to Livestock Production*, Montpellier. **[AU5: Please provide name and location of proceedings publisher.]**
- Jensen, J. 2001. Genetic evaluation of dairy cattle using test-day models. *J. Dairy Sci.* 84:2803–2812.
- López-Romero, P., and M. J. Carabaño. 2003. Comparing alternative random regression models to analyse first lactation daily milk yield data in Holstein-Friesian cattle. *Livest. Prod. Sci.* 82:81–96.
- López-Romero, P., R. Rekaya, and M. J. Carabaño. 2003. Assessment of homogeneity vs heterogeneity of residual variance in random regression test-day models in a Bayesian analysis. *J. Dairy Sci.* 83:3374–3385.
- Meyer, K. 2002. “RRGIBBS” – A program for simple random regression analyses via Gibbs Sampling. *Proc. 7th World Congress of Genetics Applied to Livestock Production*, Montpellier. **[AU6: Please provide name and location of proceedings publisher.]**
- Meyer, K. 2005. Estimates of genetic covariance function for growth of Angus cattle. *J. Anim. Breed. Genet.* 122:73–85.
- Misztal, I., T. Strabel, J. Jamrozik, E. A. Mantysaari, and T. H. Meuwissen. 2000. Strategies for estimating the parameters needed for different test-day models. *J. Dairy Sci.* 83:1125–1134.
- Moioli, B., and A. Borghese. 2005. Buffalo breeds and management systems. Pages 51–76 in *Proc. Buffalo Production Research*. Istituto Sperimentale per la Zootecnia. Animal, Monterotondo, Rome, Italy. **[AU7: Please provide name and location of proceedings publisher.]**
- Olori, V. E., W. G. Hill, S. McGuirk, and S. Brotherstone. 1999. Estimating variance components for test day milk records by restricted maximum likelihood with random regression animal model. *Livest. Prod. Sci.* 61:53–63.
- Peeva, T. 2002. Genetic improvement of buffaloes in Bulgaria. *Proc. 1st Bufalo Symposium of Americas*, Belém, Pará, Brazil. **[AU8: Please provide name and location of proceedings publisher.]**
- Raftery, A. E., and S. Lewis. 1992. How many iterations in the Gibbs sampler? Pages 763–773 in *Bayesian Statistics 4*. J. M. Bernardo et al. **[AU9: Please add names of all editors]**, ed. Oxford University Press, Oxford, UK.
- Roos, A. P. W., A. G. F. Harbers, and G. Jong. 2004. Random herd curves in a test-day model for milk, fat, and protein production of dairy cattle in the Netherlands. *J. Dairy Sci.* 87:2693–2701.
- Rosati, A., and L. D. Van Vleck. 1998. Estimation of genetic parameters for milk, fat, protein and mozzarella cheese production in the Italian river buffalo population. *Proc. 6th World Congress on Genetic Applied to Livestock Production*, Armidale, Australia. **[AU10: Please provide name and location of proceedings publisher.]**
- Smith, B. J. 1997. BOA: An R Package for MCMC Output Convergence Assessment and Posterior Inference. *J. Stat. Softw.* 21:2–37.
- Tonhati, H., M. F. C. Muñoz, J. A. Oliveira, J. M. C. Duarte, T. P. Furtado, and S. P. Tseimazides. 2000. Parâmetros genéticos para a produção de leite, gordura e proteína em bubalinos. *R. Bras. Zootec.* 29:2051–2056.