Multiple-trait random regression models for the estimation of genetic parameters for milk, fat, and protein yield in buffaloes

Rusbel Raul Aspilcueta Borquis,* Francisco Ribeiro de Araujo Neto,* Fernando Baldi,* Naudin Hurtado-Lugo,* Gregório M. F. de Camargo,* Milthon Muñoz-Berrocal,† and Humberto Tonhati*‡[1]
*Department of Animal Science, São Paulo State University, Jaboticabal, SP, Brazil, 14884 900
†Universidad Nacional Agraria de la Selva, Aptdo. 156, UNAS-Tingo Maria-Peru
‡Conselho Nacional de Desenvolvimento Científico e Tecnologico and Instituto Nacional de Ciência e Tecnologia – Ciência Animal, Viçosa, MG, Brazil, 36570 000

ABSTRACT

In this study, genetic parameters for test-day milk, fat, and protein yield were estimated for the first lactation. The data analyzed consisted of 1,433 first lactations of Murrah buffaloes, daughters of 113 sires from 12 herds in the state of São Paulo, Brazil, with calvings from 1985 to 2007. Ten-month classes of lactation days were considered for the test-day yields. The (co)variance components for the 3 traits were estimated using the regression analyses by Bayesian inference applying an animal model by Gibbs sampling. The contemporary groups were defined as herd-year-month of the test day. In the model, the random effects were additive genetic, permanent environment, and residual. The fixed effects were contemporary group and number of milkings (1 or 2), the linear and quadratic effects of the covariable age of the buffalo at calving, as well as the mean lactation curve of the population, which was modeled by orthogonal Legendre polynomials of fourth order. The random effects for the traits studied were modeled by Legendre polynomials of third and fourth order for additive genetic and permanent environment, respectively, the residual variances were modeled considering 4 residual classes. The heritability estimates for the traits were moderate (from 0.21–0.38), with higher estimates in the intermediate lactation phase. The genetic correlation estimates within and among the traits varied from 0.05 to 0.99. The results indicate that the selection for any trait test day will result in an indirect genetic gain for milk, fat, and protein yield in all periods of the lactation curve. The accuracy associated with estimated breeding values obtained using multi-trait random regression was slightly higher (around 8%) compared with single-trait random regression. This difference may be because to the greater amount of information available per animal.

Key words: covariance functions, heritability, Legendre polynomials

INTRODUCTION

Buffalo milk is an important source of nutrients because it contains higher percentages of protein and fat than cattle milk, with a fat content varying from 6.87 to 8.59% and a protein content varying from 3.91 to 4.55% (Rosati and Van Vleck, 2002; Aspilcueta-Borquis et al., 2010a). This richness confers particularities to the production, yield, and quality of products derived from fresh buffalo milk. Currently, the dairy industry favors milk containing high levels of solids, as most of milk is processed into dairy products (yogurt, cheese, and so on), a fact that makes buffalo milk a profitable alternative.

Milk yield records and their components measured along the lactation provide a typical example of repeated measurements or longitudinal data (measurements obtained sequentially from the same animal over time). The main feature of longitudinal data is the evaluation of the genetic properties of one or more variables across a specific dimension, such as time or space (Wiggans and Goddard, 1997). Several studies showed that random regression models (**RRM**) are an appealing approach to model repeated measurements and to estimate genetic parameters for these traits (Schaeffer and Jamrozik, 2008; Sesana et al., 2010).

In contrast to multi-trait models (finite-dimensional models) that provide point predictions, RRM permit the prediction of breeding values for the lactation curve as a whole, for any point or interval within the lactation period, and for functions of the lactation curve (Jamrozik et al., 1997). In dairy cattle, the regressions are fitted as a function of the lactation period using a base function that permits modeling of the trajectory for both the population mean (fixed regression) and each animal (random regression). Another advantage of RRM is that the covariance structure between measures

Table 1. Descriptive statistics of the test-day data of milk, fat, and protein yield

Item	Milk yield	Fat yield	Protein yield
Observations (n)	10,983	10,691	10,691
Sires (n)	113	113	113
Cows (n)	1,433	1,433	1,433
Mean (kg)	6.08	0.38	0.24
SD (kg)	2.32	0.14	0.09
Minimum (kg)	1.0	0.07	0.05
Maximum (kg)	14.40	0.87	0.54
CG[1]	708	676	676

[1]CG = contemporary group.

can be characterized by a function, a fact that permits greater flexibility in the use of longitudinal data (Van Der Werf and Schaeffer, 1997).

Although, several studies report genetic parameter estimates for test-day milk yield records and their components obtained by RRM in dairy cattle (Schaeffer and Jamrozik, 2008) and buffaloes (Breda et al., 2010; Sesana et al., 2010; Aspilcueta-Borquis et al., 2010b), few studies apply multi-trait random regression for test-day milk, fat, and protein yield (Muir et al., 2007; Hammani et al., 2008). The implementation of a multi-trait random regression in dairy is attractive because it allows for the study of the genetic relationships between milk volume and milk components along the lactation. According to Hammani et al. (2008), the multiple-trait RRM for simultaneous analysis of all yields seems to be a logical choice. In buffaloes, this issue is very important, as buffalo milk is valued by constituents and the lactation curve in buffaloes is different from that observed in cattle, from a biological point of view. Compared with dairy cattle, total milk yield of buffaloes is lower, the lactation length is shorter, and the lactation peak is less accentuated. In view of the possibility to use RRM and of the importance of milk components for buffalo farming, studies investigating the association between test-day milk, protein, and fat yield along the lactation are needed. Therefore, the objective of the present study was to estimate (co)variance components for first lactation milk, fat, and protein yield by single and multi-trait RRM using Bayesian methods.

MATERIALS AND METHODS

In the present study, milk, fat, and protein yield test-day records from 1,433 first lactation Murrah buffaloes, daughters of 113 sires, were analyzed. The buffaloes had calved between 1985 and 2007 and belong to 12 herds from the state of São Paulo, Brazil. The milking test controls were performed monthly, keeping the calves at dams to stimulate milk production.

The milk yield records were obtained starting on the fifth day after calving and were truncated at 305 d of lactation, as only 12% of females had a lactation length greater than this period. Only buffaloes that had their first test-day record before 45 d after calving were considered in the analyses. The test-day productions were considered in monthly lactation classes, varying from 1 to 10 classes, and included animals with at least 4 tests. The contemporary groups were defined as herd-year-month of milk test, with the restriction that each group had to contain at least 4 animals. After data editing, the descriptive statistics for milk, fat, and protein yield along the lactation (first lactation) are shown in Table 1. A pedigree file containing 2,092 animals was used in all the analyses.

Single-trait and 3-trait RRM (monthly test-day for milk, fat, and protein yield) were used for the analysis. The model included contemporary group and number of milkings per day (1 or 2 levels) as fixed effects, age of buffalo at calving (linear and quadratic effects) as covariate, and the average lactation curve of the population, which was modeled using a fourth-order orthogonal polynomial. The additive genetic, permanent environmental, and residual effects were included as random effects. To model additive genetic and permanent environmental random effects for all traits, third- and fourth-order Legendre polynomials were used. Residual variances were modeled using a step function with 4 classes as follows: 1, 2 to 3, 4 to 8, and 9 to 10 mo of lactation. The order of polynomials and residual classes applied were defined based on preliminary single-trait RRM analyses.

The (co)variance components were estimated by Bayesian inference using the Gibbs sampler of the GIBBS3F90 program (Misztal, 2010). The matrix representation of the model can be described as

$$y = X\beta + Za + Wp + e,$$

where y is the vector of phenotypic records (milk, fat, and protein yield); β, a, p, and e are vectors of fixed effects, additive genetic coefficients, permanent environmental coefficients, and residual coefficients, respectively; and X, Z, and W are incidence matrices relating