Buffalos milk yield analysis using random regression models

C.V. de Araújo¹, A. Amorim Ramos², S. Inoe Araújo¹, L. Celi Chaves¹, A.S. Schierholt³

¹Universidade Federal Rural da Amazônia – Belém – Pará, Brazil
²Universidade Estadual Paulista – UNESP – Botucatu – São Paulo, Brazil
³Universidade Federal de Viçosa – Viçosa – Minas Gerais, Brazil

Corresponding author: C.V. de Araújo. Universidade Federal Rural da Amazônia - Tel. 091 3210 5223 - Email: araujocv@bol.com.br

ABSTRACT: Data comprising 1,719 milk yield records from 357 females (predominantly Murrah breed), daughters of 110 sires, with births from 1974 to 2004, obtained from the Programa de Melhoramento Genético de Bubalinos (PROMEBUL) and from records of EMBRAPA Amazônia Oriental - EAO herd, located in Belém, Pará, Brazil, were used to compare random regression models for estimating variance components and predicting breeding values of the sires. The data were analyzed by different models using the Legendre’s polynomial functions from second to fourth orders. The random regression models included the effects of herd-year, month of parity date of the control; regression coefficients for age of females (in order to describe the fixed part of the lactation curve) and random regression coefficients related to the direct genetic and permanent environment effects. The comparisons among the models were based on the Akaike Information Criterion. The random effects regression model using third order Legendre’s polynomials with four classes of the environmental effect were the one that best described the additive genetic variation in milk yield. The heritability estimates varied from 0.08 to 0.40. The genetic correlation between milk yields in younger ages was close to the unit, but in older ages it was low.

Key words: Genetic evaluation, Milk yield, Murrah Breed, Random regression.

INTRODUCTION - The dairy production in Brazil is coming through a restructuring of it’s production system, that is aimed mainly on increasing efficiency and competitiveness. In all the country regions of Brazil, buffalos constitute a good economic option, mainly through dairy production, leading to a development of this sector. But it is necessary that a great number of studies evaluate new alternatives for a better dairy production, highlighting the genetics, investment, and breeding animals more adaptable to the environment with higher production capabilities. Genetic improvement is done by using genetically proven breeders. The genetic evaluations have been adopting models that contain daily production data - the most recent methodology is the random regression models. These models allow breeding value prediction for the lactation curve at any desired point on the time scale used, without the need of creating arbitrary performance classes or the use of adjustment factors for certain ages, al-
ollowing the prediction of the trait values in any point among this interval. The objective of this paper was to identify the best random regression model related to parameter numbers and residual classes for the characterization of genetic variation in dairy buffalo's production.

**MATERIAL AND METHODS** - We used 1,719 records of milk yield means in lactation of 357 females buffalos, mainly from Murrah breed, daughters of 110 sires, in 12 herds located in the north, northeast and southeast of Brazil, with the calvings occurring from 1974 to 2004. The registers were obtained from the Programa de Melhoramento Genético de Bubalinos (PROMEBUL) and also from records of EMBRAPA Amazônia Oriental - EAO herd, located in Belém, Pará, Brazil.

The milk yield average records in lactation were analyzed, for the covariance components and genetic parameter estimations, by random regression models. For the shape of the lactation curves descriptions, we used the Legendre's polynomial function models, by obtaining the random regression coefficients related to the fixed and random components of the model. The models used were different among them by their polynomials adjustment order (k), which varied from second to fourth order (k = 3, 4 and 5). The model included the fixed effects of levels of herd-year and month and random regression coefficients of the random additive genetic effect and the permanent environment effect for each individual. With the objective aimed to verify presence of variance heterogeneity for the temporary environmental effect, this one was considered constant along the whole lactation curve in the random regression models whose Legendre’s polynomials adjustment was of second order, and lately, compared in situations in which this same effect was allocated in two, three, four and five residual classes. The random regression models were compared among each other by the Akaike Information Criterion (AIC).

**RESULTS AND CONCLUSIONS** - The values of the likelihood function logarithms for the AIC in models with different classes of temporary environmental effects and different adjustment levels of the covariance function are shown in Table 1.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>K3E1</td>
</tr>
<tr>
<td>Log(L)</td>
<td>-1406,49</td>
</tr>
<tr>
<td>AIC</td>
<td>2838,98</td>
</tr>
</tbody>
</table>

The models that used adjustment of second order were more efficient as the number of residual classes were higher. Although the polynomial model of fourth order with four residual classes has shown the lowest value for the likelihood function logarithms among all the models, the model of third order with four residual classes has shown the lowest value for the AIC, thus
being the most efficient to describe the milk yield variation along the lactation curve. Independently of the residual class numbers, the models that used Legendre’s polynomial of second order, described the variances in a similar way, with a variance overestimation at the end of the curve for models with higher residual classes. For models that considered four and five parameters (third and fourth order, respectively), the description of the additive genetic variance behavior was similar to the other models, but, nearly after 132 months of age at calving, the models of second order had a variance overestimation, mainly at the end of the lactation curve. In this way, for the description of the additive genetic variation of milk yield along the lactation curve that describes the female age at calving, random regression models must apply adjustment orders higher than the second.

By considering the temporary environment effect constant along the curve that describes the female age at calving, when this one was heterogenic, it caused variance overestimation in the beginning of the curve and sub estimation in the middle. It was verified that models with polynomial adjustments from second to fourth order, adopting four residual classes, described the variation of this effect in a similar way along the curve. The heritability estimates obtained in each model with different polynomial adjustment order and different numbers of temporary effect classes are shown in Figure 1.

Figure 1. Herdability estimates for milk yield, along the curve, for models of second order with one, two, three, four and five residual classes K3E1, K3E2, K3E3, K3E4 e K3E5, respectively, and for models of third and fourth order with four residual classes, K4E4 E K5E4, respectively.

The formation of different residual classes for the temporary environmental effect did not cause great alterations in the herdability estimates in the models with an adjustment of second order. The comparison of estimates among models with different polynomial adjustments has shown that models of third and fourth order had very similar behavior, in a way that estimates in a more advanced age were lower than those predicted with models considering less parameters.

The milk yield genetic correlations were high and positive when ages were next to each other. On the other hand, between ages belonging to more extremes points at the curve that describes age, the correlation was negative. The negative genetics correlations between the-
se extremes points at the curve shows that selection for a higher milk yield in certain lactation age won’t have a positive effect on milk yield in females with more advanced age. In studies that used milk yield on test day control, in first lactation dairy cattle, Araújo (2003), Brotherstone et al. (2000), Cobuci (2002) and Costa et al. (2002) verified a lower negative genetic correlation estimates between these extremes points in the lactation curve.

The random regression models were efficient to describe the milk yield genetic variation, being more efficient than those that used the Legendre’s polynomial function of fourth order, admitting four residual classes, minimizing the residual variance, when compared with the others models with different adjustment orders and different numbers of residual classes.