

Principal Components for Reproductive and Productive Traits in Buffaloes from Brazil

D.P. Oliveira^{1,2}, C.C. Barros¹, F.R. Araujo Neto¹, D.A.L. Lourenco³, N.A. Hurtado-Lugo¹ and H. Tonhati¹

¹Universidade Estadual Paulista “Júlio de Mesquita Filho” (FCAV-UNESP), Jaboticabal, São Paulo, Brazil. ²CAPES Foundation, Ministry of Education of Brazil, Brasília, Distrito Federal, Brazil. ³University of Georgia, Athens, GA, USA.

ABSTRACT: A total of 12,450 records for first-parity in dairy buffaloes were used in this study. Analyzed traits were age at first calving, interval between first and second calving, lactation length, somatic cell count, and milk, fat, protein, lactose, and total solids yields. The covariance components were estimated through principal components models by reduced-rank with third to ninth components. Direct additive genetic and residual random effects were included in all models. The three first eigenvalues explained 94.25% of the genetic variance on average. The estimates of heritability ranged from 0.03 to 0.31. The genetic correlation of production traits had moderate to high magnitude, and reproductive traits had negative and low to moderate magnitude. According to the selection criterion, the four genetic principal components are appropriate for evaluation of economic traits for Buffaloes in Brazil.

Keywords: Dairy buffaloes; Reduced-rank; Economic traits

Introduction

There are about 195 million buffaloes worldwide. According to FAO (2011) the majority of these animals are in Asia, but there are almost 1.2 million animals in Brazil. Buffaloes provide meat, milk and work ability; therefore, they are considered as a species with social and economic importance mainly in the developing countries. However, buffaloes have been a good economic alternative worldwide due to milk yield and the production of *mozzarella* cheese (Tonhati et al. (2000)).

The genetic analyses for reproduction and milk production from buffaloes used multi-trait models to study the association among several traits and determine an appropriate selection criteria for the population. However, large data sets and large number of traits complicate the analysis due to the time demands and the high computational requirements.

The genetic evaluation of a major number of traits has been performed with two or three traits to estimate the total variance (Meyer, 2005). A way of decreasing the complexity of the multi-trait analyses is through the reduction of estimated parameters and the ranking of the genetic covariance matrix. One of the tools used is genetic

principal components “PC” analysis (Boligon et al. (2013); Bignardi et al. (2012); Meyer (2007a)). The PC analysis consists of the transformation of the dataset of variable correlation into a new data set, wherein are linearly combined from the original uncorrelated data that explains the maximum variance (Meyer and Kirkpatrick (2005)).

This study aims to compare genetic PC models for economic traits of the Murrah Buffaloes in Brazil.

Material and methods

A total of 12,450 records from 2,531 first-parity lactations of Murrah Buffaloes (*Bubalus bubalis*) were analyzed. Animals were 23 to 64 months old and they calved between 2005 and 2013. They belonged to 12 herds from Sao Paulo State, Brazil. The traits analyzed were age at first calving “AFC”; interval between at first and second calving “IBC”; lactation length “LL”; somatic cell count “SCC”; milk yield at 270 days “MY”; and fat “FY”, protein “PY”, lactose “LY” and total solids “TS” yields.

Milk yield from the first control was obtained from the 5th through the 45th day after calving. The contemporary groups “CG”, which contained at least three animals, were set according to herd, birth year and season for the AFC trait, and herd, calving year and season for all other traits. A pedigree with 14,638 animals was used for all analyses. The dataset structure and descriptive statistics of each studied trait are shown in Table 1.

Covariance components were estimated by the Restricted Maximum Likelihood method and the genetic PC model with reduced-rank for the third through ninth components using the Wombat program (Meyer (2007b)). The model can be represented by the following equation:

$$y = X\beta + Z^*a^* + e$$

where **y** is the vector of the observed traits; **X** is the incidence matrix of fixed effects; **β** is the vector of fixed effects (CG, milking frequency and buffalo calving age as linear and quadratic covariates); **Z*** is the incidence matrix of random additive genetic effects, where $Z^* = Z(E \otimes I)$; **a*** is the vector of random additive genetic effects, where $a^* =$

Table 1. Number of observations “N”, mean, standard deviation “SD” and number of contemporary groups “CG” for reproductive and productive traits in Murrah Buffaloes.

Traits ^κ	N	Mean	SD	GC
MY, kg	2531	1699.95	572.58	138
LL, days	2531	269.57	65.12	138
PY, kg	863	120.07	36.56	40
PY, kg	865	77.57	23.69	40
LY, kg	628	86.79	25.53	30
TS, kg	532	289.44	85.60	26
AFC, days	2436	1140.05	171.83	138
IBC, days	1408	438.55	65.21	122
SCC	656	2.19	0.71	40

^κMY: Milk Yield, LL: Lactation length, FY: fat yield, PY: protein yield, LY: Lactose yield, TS: total solids yield, AFC: age at first calving, IBC: Interval between at first and second calving, SCC: somatic cell count.

Table 2. Number of parameters “np”, maximum log likelihood (log L), Akaike “AIC” and Bayesian “BIC” Information Criteria and proportion of direct additive genetic variance (% λ) explained by the first three eigenvalues in different genetic principal component “PC” analyses for reproductive and productive traits in Murrah Buffaloes.

Model [§]	np	log L [*]	AIC [*]	BIC [*]	% λ_1	% λ_2	% λ_3
PC3	69	207.60	276.60	530.82	95.47	4.21	0.32
PC4	75	193.61	268.61	544.93	94.15	5.29	0.42
PC5	80	189.58	269.58	564.32	94.06	5.27	0.44
PC6	84	188.70	272.70	582.19	94.03	5.23	0.45
PC7	87	188.62	275.62	596.15	94.02	5.23	0.46
PC8	89	188.66	277.66	605.57	94.02	5.23	0.46
PC9	90	188.63	278.63	610.21	94.02	5.23	0.46

[§]PC3= Three PC; PC4= Four PC; PC5= Five PC; PC6 = Six PC; PC7 = Seven PC; PC8 = Eight PC; PC9 = Nine PC.

^{*}Values of logL, AIC, BIC=-51,000.

$(E' \otimes I)$, $\text{Var}(a^*) = (I \otimes A)$ and $\Sigma a = E \Lambda E'$. E is a matrix of eigenvectors e_i . Λ = diagonal matrix of eigenvalues λ_i , and \otimes is the direct product. It is assumed that λ_i and e_i are in decreasing order of the magnitude of λ_i . Thus, the matrix of genetic covariance (Σ) can be decomposed in term of E and Λ , with $E E' = I$. To consider m principal components, the matrix E is replaced by E_m , which comprises m first columns of E , therefore, E_m is given to define Z^* and a^* . The number of equations is reduced, resulting in reduction of Λ . The residual (co)variance matrix was assumed to have full rank (Kirkpatrick and Meyer (2005)).

Models were compared by Bayesian “BIC” and Akaike “AIC” Information Criterion. These criteria allow the comparison of non-hierarchical models and penalizes models with higher number of parameters.

Results and Discussion

For the studied models, the Maximum Likelihood Logarithm increased with the increment of the number of parameters in the model. However, up to PC6 models, this increase was slight. The AIC had an increase in value with the increment of the number of parameters, except for PC3 (276.60). The AIC indicated that PC4 (268.61) had the best fit. But the BIC indicated that PC3 had the best fit (Table 2).

The genetic variance explained by the first three eigenvalues was similar to the increase of the number of model parameters up to PC4. Production traits were more strongly correlated among themselves and more weakly correlated with the reproductive traits. This might explained the slight difference in the genetic variance explained by the first three eigenvalues (Bignardi et al. (2012)). On the other hand, Boligon et al. (2013) observed contrary results for performance traits in the Nellore cattle.

The heritability for productive traits ranged from 0.03 to 0.30, for AFC (0.14) and IBC (0.03) in PC3. The heritability for productive traits ranged from 0.17 to 0.30, for AFC (0.19) and IBC (0.04) in PC4. The heritabilities were similar between the models, except to SCC (PC3=0.03; PC4=0.17). However, the heritabilities for SCC in PC4 were similar to values found in the literature (Table 3). Malhado et al. (2013); Rodrigues et al. (2010); Aspilcueta-Borquis et al. (2010) observed similar heritability for the same traits.

The heritabilities were similar for all traits from PC4 to PC9, except for SCC and IBC. SCC and IBC showed an increase of 35.6 % and 50 %, respectively, compared to the full rank model (PC9). These traits have low genetic association with the others, which indicated that the model with reduced rank was more efficient when the traits had a higher association among themselves.

The production traits MY, LL, FY, FP, PY, LY, TS showed positive genetic correlation among themselves in PC4. The values were of moderate to high magnitude and ranged from 0.58 to 0.99. The AFC and IBC had negative genetic association of low to moderate magnitude when combined with MY, LL, FY, FP, TS. The values ranged from -0.46 to 0.01, except to AFC with LL (0.16). The SCC showed low and negative genetic correlations between MY (-0.18), FY (-0.36), PY (-0.18), LY (-0.16), and TS (-0.26). The genetic correlation between SCC and LL was (0.43), and SCC with AFC (0.53), and SCC with IBC (0.08). These results were similar with the literature (Aspilcueta-Borquis et al. (2010)).

The genetic correlation of SCC with production and quality milk was low and negative. There were moderate to high genetic associations of SCC with LL, SCC with AFC, and SCC with IBC, which indicated that these traits are more influenced by environmental conditions. The heritability indicated that selection for MY for Dairy Buffaloes would result in genetic improvement in FY, FP, FL, TS and LL. On the other hand, the genetic correlations of MY with SCC, AFC, and IBC were of low to moderate magnitude and negative showing that the selection for MY would allow favorable gain in IBC. But a low correlation would be expected in SCC and AFC.

Thus, a subset of PC was sufficient to perform the genetic evaluation of this dataset. The results obtained for AIC and BIC criteria, the variance explained by the first three eigenvalues, the heritability and genetic correlations suggest that the PC4 was suitable to explain the genetic variance for economic traits in dairy Buffaloes in Brazil.

Conclusions

Principal components analyses are used in multi-trait models for reducing the number of model parameters without decreasing the efficiency of estimation. Fitting the four genetic principal components can be ideal to explain

the covariance structure for economic traits in Murrah Buffaloes in Brazil.

Acknowledgements

Support from Sao Paulo Research Foundation “FAPESP-Grants: 2009/53773-1” and National Council for Scientific and Technological Development “CNPq” (Brazil) are gratefully acknowledged.

Literature Cited

Aspilcueta-Borquis, R.R., Araujo Neto, F. R., Baldi, F. et al. (2010). *J. Dairy Sci.* 93:2195-2201.
 Bignardi, A.B., Faro, L.El, Rosa, G.J.M. et al. (2012). *J. Dairy Sci.* 95: 2157-2164.
 Boligon, A.A., Bignardi, A.B., Mercadante, M.E.Z. et al. (2013). *Liv. Sci.* 152:135-142.

FAO STAT (2011). <http://faostat3.fao.org/faostat-gateway/go/to/download/Q/QA/E>. Accessed on Jan. 2014.
 Malhado, C.H.M., Malhado A. C. M., Ramos A. A. et al. (2013). *R. Bras. Zootec.* 42:565-569.
 Meyer, K. (2005). *Anim. Sci.* 81: 337-345.
 Meyer, K. (2007a). *J. Anim. Breed. Genet.* 124:50–64.
 Meyer, K. (2007b). *J. Zhejiang Univ. Sci. B.* 8:815–821.
 Meyer, K., and M. Kirkpatrick. (2005). *Genet. Sel. Evol.* 37:1–30.
 Rodrigues, A. E., Marques, J. F. R., Camargo Júnior, R. N. C. et al. (2010). *Arq. Bras. Med. Vet. Zootec.* 62:712-717.
 Tonhati, H., Céron Muñoz, M. F., Oliveira, J. A. et al. (2000). *Rev. Bras. Zootec.* 6:2051-2056.

Table 3. Estimates of heritability in different genetic principal component (PC) analyses for reproductive and productive traits in Murrah Buffaloes.

Trait ^{&}	Genetic Principal Components [§]						
	PC3	PC4	PC5	PC6	PC7	PC8	PC9
MY, kg	0.30	0.30	0.30	0.30	0.30	0.30	0.30
LL, days	0.06	0.07	0.08	0.09	0.09	0.09	0.09
PY, kg	0.20	0.24	0.25	0.25	0.25	0.25	0.25
PY, kg	0.28	0.30	0.31	0.31	0.31	0.31	0.31
LY, kg	0.29	0.29	0.30	0.30	0.30	0.30	0.30
TS, kg	0.24	0.27	0.27	0.28	0.28	0.28	0.28
AFC, days	0.14	0.19	0.19	0.18	0.19	0.19	0.19
IBC, days	0.03	0.04	0.05	0.06	0.06	0.06	0.06
SCC	0.03	0.17	0.18	0.23	0.23	0.22	0.23

[§]PC3= Three PC; PC4= Four PC; PC5= Five PC; PC6 = Six PC; PC7 = Seven PC; PC8 = Eight PC; PC9 = Nine PC.

[&]MY: Milk Yield, LL: Lactation length, FY: fat yield, PY: protein yield, LY: Lactose yield, TS: total solids yield, AFC: age at first calving, IBC: Interval between at first and second calving, SCC: somatic cell count.