

Random regression models for milk, fat and protein in Colombian Buffaloes

Modelos de regresión aleatoria para leche, grasa y proteína en búfalos de Colombia

Naudin Hurtado-Lugo,^{1,2} Ph.D, Humberto Tonhati,¹ Ph.D, Raul Aspilcuelta-Borquis,¹ Ph.D, Cruz Enríquez-Valencia,¹ M.Sc, Mario Cerón-Muñoz,^{2*} Ph.D.

¹University of São Paulo, Faculty of Agrarian and Veterinary Sciences (FCAV), State Jaboticabal (UNESP), 14884900, SP, Brazil. ²Universidad de Antioquia, Facultad de Ciencias Agrarias, Grupo de investigación GaMMA, Carrera 75 No 65-87. Medellín, Colombia. Corresponding: cerongamma@gmail.com

Received: June 2013; Accepted: August 2014.

ABSTRACT

Objective. Covariance functions for additive genetic and permanent environmental effects and, subsequently, genetic parameters for test-day milk (MY), fat (FY) protein (PY) yields and mozzarella cheese (MP) in buffaloes from Colombia were estimate by using Random regression models (RRM) with Legendre polynomials (LP). **Materials and Methods.** Test-day records of MY, FY, PY and MP from 1884 first lactations of buffalo cows from 228 sires were analyzed. The animals belonged to 14 herds in Colombia between 1995 and 2011. Ten monthly classes of days in milk were considered for test-day yields. The contemporary groups were defined as herd-year-month of milk test-day. Random additive genetic, permanent environmental and residual effects were included in the model. Fixed effects included the contemporary group, linear and quadratic effects of age at calving, and the average lactation curve of the population, which was modeled by third-order LP. Random additive genetic and permanent environmental effects were estimated by RRM using third- to- sixth-order LP. Residual variances were modeled using homogeneous and heterogeneous structures. **Results.** The heritabilities for MY, FY, PY and MP ranged from 0.38 to 0.05, 0.67 to 0.11, 0.50 to 0.07 and 0.50 to 0.11, respectively. **Conclusions.** In general, the RRM are adequate to describe the genetic variation in test-day of MY, FY, PY and MP in Colombian buffaloes.

Key words: Cattle, genetics, zootechnics (*Source: EuroVoc*).

RESUMEN

Objetivo. Fueron estimadas funciones de covariancia para los efectos genéticos aditivos y de ambiente permanente, y posteriormente fueron estimados los parámetros genéticos para la producción de leche en el día del control (PDC) para leche (MY), grasa (FY), proteína (PY) y producción de queso mozzarella (MP), usando modelos de regresión aleatoria (RRM) mediante Polinomios Ortogonales de Legendre (LP) en búfalos de Colombia. **Materiales y Métodos.** Fueron analizados 1884 registros de PDC de primeras lactancias para MY, FY, PY y MP, hijas de 228 reproductores. Los registros procedían de 14 rebaños entre 1995 y 2011. Para las PDC fueron consideradas 10 clases mensuales de días

en lactancia. Los grupos contemporáneos fueron definidos como rebaño-año-mes de control. Fueron incluidos en el modelo efectos aleatorios genético aditivo, de ambiente permanente y residual. Los efectos fijos incluidos fueron grupo contemporáneo, efecto lineal y cuadrático de edad al parto, y curva promedia de la lactancia de la población, la cual fue modelada por un LP de tercera orden. Los efectos aleatorios genético aditivo y de ambiente permanente fueron estimados por RRM mediante LP de tercero a sexto orden. Las variancias residuales fueron modeladas usando estructuras homogéneas y heterogéneas. **Resultados.** Las heredabilidades para MY, FY, PY y MP variaron de 0.38 a 0.05, 0.67 a 0.11, 0.50 a 0.07 y 0.50 a 0.11, respectivamente. **Conclusiones.** En general, los RRM son adecuados para describir la variación genética para la PDC de MY, FY, PF y MP en búfalos de Colombia.

Palabras clave: bovino, genética, zootecnia (*Fuente: EuroVoc*).

INTRODUCTION

Buffalo milk is mainly used for the production of milk derivatives due to the composition and physicochemical quality of its components (1). The nutritional quality and economic value of mozzarella cheese does not only depend on the quantity of milk produced, but also on its relationship with fat and protein (1,2).

Random regression models (RRM) can be applied to test-day milk yield as an alternative to standard procedures used for the genetic evaluation of longitudinal traits in dairy cattle. According to El Faro et al (3), in dairy cattle of different breeds, heritabilities of 0.15 to 0.48, 0.13 to 0.28 and 0.16 to 0.28 were estimated for milk (MY), fat (FY) and protein (PY) yield, respectively, using RRM (4-7).

Studies estimating heritabilities for mozzarella cheese production (MP) with RRM are scarce. The objective of this study was to estimate covariance functions for additive genetic and permanent environmental effects and, subsequently, genetic parameters for MY, FY, PY and MP in buffaloes from Colombia using RRM with Legendre polynomials (LP).

MATERIALS AND METHODS

Description of the data. In the study, test-day MY, FY, PY and MP from 1884 first lactations of buffalo cows aged 26 to 49 months, from 228 sires, were analyzed. The animals belonged to 14 herds that comprise the database of Faculty of Agricultural Sciences, University of Antioquia, Colombia and Colombian Buffaloes Breeders Association. Calvings occurred between 1995 and 2011. Lactations were considered when the first control was between 5 and 45 d. Test-day yields were divided into monthly classes of lactation (1 to 10 classes) and animals with at least three test-day records were included in the analysis. The contemporary groups (CG) were defined as herd-year-month of milk test, with the restriction that each group should contain at least three animals. A pedigree file containing 7.090 animals was used in all analyses.

INTRODUCCIÓN

La leche de búfalo se utiliza principalmente para la producción de derivados lácteos, debido a la composición y calidad fisicoquímica de sus componentes (1). La calidad nutricional y el valor económico del queso mozzarella no depende únicamente de la cantidad de leche producida, sino también de su relación con la grasa y la proteína (1,2).

Es posible aplicar modelos de regresión aleatoria (MRA) a la producción de leche del día de control como una alternativa a los procedimientos estándar utilizados para la evaluación genética de características longitudinales en bovinos lecheros. De acuerdo con El Faro et al (3), en los bovinos lecheros de diferentes razas, se estimaron las heredabilidades de 0.15 a 0.48, de 0.13 a 0.28 y de 0.16 a 0.28 para la producción de leche (MY), grasa (FY) y proteína (PY), respectivamente, utilizando los MRA (4-7).

Las estimaciones obtenidas en estudios de las heredabilidades para la producción de queso mozzarella (MP) via MRA son escasas. El objetivo del presente estudio fue estimar las funciones de covarianza para los efectos genéticos aditivos y de ambiente permanente y, posteriormente, los parámetros genéticos para MY, FY, PY y MP en búfalos de Colombia utilizando MRA con polinomios de Legendre (LP).

MATERIALES Y MÉTODOS

Descripción de los datos. En el estudio fueron analizados los registros productivos de MY, FY, PY y MP en el día de control de primeras lactancias de 1.884 búfalas de 26 a 49 meses de edad, hijas de 228 reproductores. Los animales pertenecían a 14 rebaños que conforman la base de datos de la Facultad de Ciencias Agrarias, Universidad de Antioquia, Colombia y la Asociación Colombiana de Criadores de Búfalos. Los partos ocurrieron entre el año 1995 y el 2011. Fueron consideradas lactancias cuando su primer control estuvo entre el día 5 y 45. La producción en el día de control se dividió en clases mensuales de lactancia (clases de 1 a 10) y se incluyeron en el análisis los animales con al menos tres registros de día de control. Los grupos contemporáneos (CG) se

MP was estimated using the mozzarella index proposed by Borghese and Mazzi (8), which is based on MY, %FY and %PY. This index is adopted in genetic and animal breeding programs in Italy and is calculated as follows:

$$MP(kg) = MY * \{ [(3.5 * \%P) + (1.23 * \%F) - 0.88] / 100 \}.$$

Statistical method. Single-trait RRM were used for the analysis of MY, FY, PY, and MP. Random additive genetic effects, permanent environmental effects of the animal, and residual effects were included in all models. The fixed effects were CG, linear and quadratic effects of the covariate age of dam at calving, and the average lactation curve of the population modeled by a third-order orthogonal polynomial. The matrix representation of the model is given as:

$$y = Xb + Za + Wap + e$$

Where

y = vector of observations;

b = vector of fixed effects (CG and age of dam at calving and average curve of the population covariates);

a = vector of solutions for additive genetic random regression coefficients;

ap = vector of solutions for permanent environmental random regression coefficients;

e = vector of residual effects, and

X, Z, W = incidence matrices for fixed, random additive genetic and permanent environmental effects, respectively.

The dimension of vector "a" consists of $k_a \times N_a$ coefficients, where k_a corresponds to the order of the polynomial and N_a is the number of animals in the relationship matrix. The dimension of vector ap consists of $k_{ap} \times N_d$ coefficients, where k_{ap} corresponds to the order of the polynomial and N_d is the number of animals with records.

For analysis, it was assumed that the records were distributed with mean $X\beta$ and the random effects were assumed to be normally distributed with:

$$\begin{bmatrix} a \\ ap \\ e \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} K_a \otimes A & 0 & 0 \\ 0 & K_{ap} \otimes I_{N_d} & 0 \\ 0 & 0 & R \end{bmatrix} \right)$$

Where

K_a and K_{ap} are covariance matrices between additive genetic and permanent environmental random regression coefficients, respectively;

A is the relationship matrix between individuals;

I_{N_d} is the identity matrix of dimension N_d ;

\otimes is the Kronecker product between matrices, and R is a block diagonal matrix containing

definieron como rebaño-año-mes de control, con la restricción que cada grupo debe contener al menos tres animales. Se utilizó un archivo del pedigrí de 7.090 animales en todos los análisis.

La MP fue estimada utilizando el índice de mozzarella propuesto por Borghese y Mazzi (8), el cual se basa en MY, %FY y %PY. Este índice ha sido adoptado en programas de genética y reproducción animal en Italia y se calcula de la siguiente forma:

$$MP(kg) = MY * \{ [(3.5 * \%P) + (1.23 * \%F) - 0.88] / 100 \}.$$

Método estadístico. Se utilizaron MRA univariados para el análisis de MY, FY, PY y MP. Se incluyeron los efectos aleatorios fueron genéticos aditivos, de ambiente permanente y los efectos residuales de los animales en todos los modelos. Los efectos fijos fueron CG, los efectos lineales y cuadráticos de la edad como covariable al momento del parto y la curva promedia de lactancia de la población, modelada por medio de un polinomio ortogonal de tercer orden. La representación matricial del modelo es la siguiente:

$$y = Xb + Za + Wap + e$$

Donde

y = vector de observaciones;

b = vector de efectos fijos (CG y las covariables edad al momento del parto y curva promedio de la población);

a = vector de soluciones para los coeficientes de regresión aleatoria genético aditivo;

ap = vector de soluciones para los coeficientes de regresión aleatoria de ambiente permanente;

e = vector de efectos residuales, y

X, Z, W = matrices de incidencia para los efectos fijos, aleatorios genético aditivo y de ambiente permanente, respectivamente.

La dimensión del vector "a" se compone de $k_a \times N_a$ coeficientes, donde k_a corresponde al orden del polinomio y N_a al número de animales en la matriz de relaciones. La dimensión del vector "ap" consiste de $k_{ap} \times N_d$ coeficientes, donde k_{ap} corresponde al orden del polinomio y N_d al número de animales con registros.

Para el análisis se asumió que los registros tenían una distribución con media $X\beta$ y se asumió que los efectos aleatorios tenían una distribución normal:

$$\begin{bmatrix} a \\ ap \\ e \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} K_a \otimes A & 0 & 0 \\ 0 & K_{ap} \otimes I_{N_d} & 0 \\ 0 & 0 & R \end{bmatrix} \right)$$

Donde

K_a y K_{ap} corresponden a las matrices de covarianza entre los coeficientes de regresión aleatorios del efecto genético aditivo y de ambiente permanente, respectivamente;

A es la matriz de relaciones entre los individuos;

I_{N_d} es la matriz identidad de dimensión N_d ;

residual variances. It was assumed that the residuals are independent.

The random additive genetic and permanent environmental effects were modeled using third- to sixth-order LP. A homogeneous (R) and a heterogeneous (r) structure with 10 classes (each month was considered to be a different class) were adopted for residual variances. In addition, heterogeneous variances with smaller classes were also considered, which were divided according to similarity between variances based on the 10 residual classes. Thus, test-day records for MY were divided into four classes (1st, 2nd to 4th, 5th to 8th and 9th to 10th month). The residual variance structure for FY was defined by four classes (1st, 2nd, 3rd to 8th and 9th to 10th month). For PY, residual variance was modeled using four classes (1st, 2nd to 7th, 8th and 9th to 10th month). Residual variance for MP was modeled using four classes (1st to 2nd, 3rd to 7th, 8th and 9th to 10th month).

The general RRM is given as $LEGk_a.k_{ap}R$ or $LEGk_a.k_{ap}r$, where k_a and k_{ap} correspond to the order of the covariance function for additive genetic and permanent environmental effects, respectively, and R or r is the structure of residual variances. The covariance functions were estimated by the restricted maximum likelihood method using the WOMBAT program (9).

The different models were compared using the logarithm of the likelihood function ($\log L$), the likelihood ratio test (LRT) at 1% probability, restricted maximum likelihood Akaike's (AIC) and Schwarz Bayesian (BIC) Information Criteria, and examination of variances and correlations estimated for the traits. The information criteria can be described as: $AIC = -2\log L + 2p$ and $BIC = -2\log L + p\log(N - r(X))$, where p is the number of parameters estimated, N is the number of data, $r(X)$ is the rank of the incidence matrix of fixed effects in the model, and $\log L$ is the logarithm of the restricted maximum likelihood function.

RESULTS

The mean test-day MY data followed a typical lactation curve for buffaloes, starting with 3.69 kg and increasing until peak yield in the third month (4.14 kg), followed by a decrease until the end of lactation (2.34 kg). The mean test-day FY, PY and MP were 0.24, 0.14 and 0.75 kg, respectively, with a standard deviation of 0.11, 0.07 and 0.33 kg and coefficient of variation of 46.58, 45.89 and 44.41% (Table 1).

⊗ es el producto de Kronecker entre las matrices, y R es una matriz diagonal que contiene las varianzas residuales. Se asume que los residuos son independientes.

Los efectos genético aditivo y de ambiente permanente fueron modelados usando LP de tercer a sexto orden. Se adoptó una estructura homogénea (R) y heterogénea (r) con 10 clases (se consideró que cada mes correspondía a una clase diferente) para varianzas residuales. Además, también se consideraron las varianzas heterogéneas con clases más pequeñas, las cuales fueron divididas según su similitud entre las varianzas en base a las 10 clases residuales. Por lo tanto, los registros del día de control para MY se dividieron en cuatro clases (1^o, 2^o a 4^o, 5^o a 8^o y 9^o a 10^o mes). La estructura de la varianza residual para FY se definió de acuerdo a cuatro clases (1^o, 2^o, 3^o a 8^o y 9^o a 10^o mes). Para PY, la varianza residual fue modelada usando cuatro clases (1^o, 2^o a 7^o, 8^o y 9^o a 10^o mes). La varianza residual para MP fue modelada usando cuatro clases (1^o a 2^o, 3^o a 7^o, 8^o y 9^o a 10^o mes).

El MRA general se define como $LEGk_a.k_{ap}R$ o $LEGk_a.k_{ap}r$, donde k_a y k_{ap} corresponden al orden de la función de covarianza para los efectos genético aditivo y de ambiente permanente, respectivamente, y R o r es la estructura de las varianzas residuales. Las funciones de covarianza fueron estimadas por el método de máxima verosimilitud restringida utilizando el programa WOMBAT (9).

Los diferentes modelos fueron comparados mediante el logaritmo de la función de verosimilitud ($\log L$), la prueba de razón de verosimilitud (LRT) con una probabilidad del 1%, los criterios de información de Akaike (AIC) y Bayesiano de Schwarz (BIC), y por medio del examen de las correlaciones y varianzas estimadas para las características. Los criterios de información pueden describirse como: $AIC = -2\log L + 2p$ y $BIC = -2\log L + p\log(N - r(X))$, donde p es el número de parámetros estimados, N es el número de datos, $r(X)$ es el rango de la matriz de incidencia de los efectos fijos en el modelo, y $\log L$ es el logaritmo de la función de máxima verosimilitud restringida.

RESULTADOS

Los datos promedios del día de control para MY siguieron una curva de lactancia típica para los búfalos, iniciando con 3.69 kg y aumentando hasta el pico de la producción en el tercer mes (4.14 kg), seguido por una disminución hasta el fin de la lactancia (2.34 kg). Los valores promedio del día de control para FY, PY y MP fueron, 0.24, 0.14 y 0.75 kg, respectivamente, con una desviación estándar de 0.11, 0.07 y 0.33 kg y con un coeficiente de variación de 46.58, 45.89 y 44.41 (Tabla 1).

Table 1. Description of milk yield (MY), fat yield (FY), protein yield (PY) and mozzarella production (MP) across first lactation of buffaloes.

Test day	N	MY			N	FY			PY			MP		
		Mean	SD	CV		Mean	SD	CV	Mean	SD	CV	Mean	SD	CV
1	711	3.69	1.33	36.1	100	0.27	0.12	44.93	0.17	0.07	39.72	0.89	0.35	39.56
2	764	4.08	1.39	34.2	118	0.27	0.13	49.21	0.17	0.08	46.68	0.90	0.41	45.77
3	808	4.14	1.34	32.2	133	0.27	0.11	40.64	0.17	0.06	39.06	0.87	0.33	38.18
4	758	4.01	1.25	31.1	141	0.24	0.10	42.06	0.15	0.05	35.06	0.78	0.29	36.59
5	749	3.82	1.14	29.7	138	0.24	0.09	38.94	0.14	0.05	37.32	0.75	0.27	36.13
6	707	3.66	1.21	32.9	123	0.23	0.11	48.28	0.14	0.06	43.26	0.73	0.32	43.16
7	705	3.37	1.14	33.8	134	0.21	0.10	47.37	0.11	0.05	48.05	0.63	0.29	46.37
8	626	3.08	1.08	35.2	88	0.18	0.08	45.23	0.10	0.04	40.98	0.54	0.22	40.70
9	536	2.78	0.97	34.99	47	0.17	0.07	42.31	0.09	0.03	34.21	0.50	0.18	35.40
10	426	2.34	0.98	41.96	14	0.15	0.07	43.03	0.08	0.03	32.35	0.46	0.17	36.31

N: number of observations; SD: standard deviation; CV (%): coefficient of variation.

Milk component yields remained constant during early lactation until the third month (0.27 kg for FY and 0.17 kg for PY). Production declined after the fourth month with increasing number of days in milk until the end of lactation (0.15 kg for FY and 0.08 kg for PY). In contrast, MP increased from early lactation (0.89 kg) to the second month (0.90 kg), followed by a decline after the third month to the end of the lactation (0.46 kg). In general, a greater variation in FY, PY and MP was observed at the end of lactation.

The LEG3,3_4 model should be used for MY, PY and MP, and the LEG3,3_6 or LEG3,3_4 models should be used for FY, with the biological interpretation of this trait being easier with the LEG3,3_4 model because of the smaller number of parameters (Table 2).

RRM with LP require the definition of the most appropriate order for each random effect included in the model. For MY, FY, PY and MP (Table 3), improvement of the log L ($p > 0.01$ by the LRT) and AIC criteria was generally observed when the order of fit was increased from three to four for additive genetic variance combined with an order of six for permanent environmental variance.

Table 4 shows the eigenvalues associated with the matrix of random regression coefficients for additive genetic and permanent environmental effects obtained for the best models for MY, FY, PY and MP according to the BIC. As can be seen, lower-order RRM permit to model most part of the additive genetic variation in MY, FY, PY and MP, with a fourth coefficient being associated with an eigenvalue of zero, in contrast to the permanent environmental effect. The first eigenvalue analyzed was responsible for at least 62% of the data variation in MRA adjusted to the traits studied.

La producción para componentes lácteos permanecieron constantes al inicio de la lactancia hasta el tercer mes (0.27 kg para FY y 0.17 kg para PY). La producción disminuyó después del cuarto mes con un mayor número de días de lactancia hasta el final de la misma (0.15 kg para FY y 0.08 kg para PY). En contraste, el MP aumentó desde al inicio de la lactancia (0.89 kg) hasta el segundo mes (0.90 kg), seguido por un descenso después del tercer mes hasta el final de la lactancia (0.46 kg). En general, se observó una mayor variación en FY, PY y MP al final del periodo de lactancia.

El modelo LEG3,3_4 fue utilizado para MY, PY y MP y el modelo LEG3,3_6 o LEG3,3_4 fueron utilizados para FY, dado que la interpretación biológica de este rasgo es más fácil con el modelo LEG3,3_4 debido al menor número de parámetros (Tabla 2).

El MRA con LP requiere de la definición del orden más apropiado para cada efecto aleatorio incluido en el modelo. Para MY, FY, PY y MP (Tabla 3), se observó de manera general una mejora en los criterios log L ($p > 0.01$ por el LRT) y AIC cuando el orden de ajuste se aumentó de tres a cuatro para la varianza genética aditiva, en combinación con un orden de seis para la varianza de ambiente permanente.

En la tabla 4, son presentados los valores propios asociados con la matriz de los coeficientes de regresión aleatoria para los efectos genético aditivo y de ambiente permanente obtenidos para los mejores modelos de MY, FY, PY y MP de acuerdo con el BIC. Como puede observarse, un MRA con un orden menor permite modelar la mayor parte de la variación genética aditiva en MY, FY, PY y MP, con un cuarto coeficiente asociado con un valor propio de cero, en contraste con el efecto de ambiente permanente. El primer valor propio analizado fue responsable de por lo menos el 62% de la variación en los datos en el MRA ajustado a los rasgos objeto de estudio.

Table 2. Number of parameters estimated (P), logarithm of the maximum likelihood function (log L), Akaike information criterion (AIC), Schwarz Bayesian Information Criterion (BIC) and Likelihood Ratio Test (LRT) for the models used to compare the number of residual variance classes for milk yield (MY), fat yield (FY), protein yield (PY), and mozzarella production (MP) in buffaloes.

Model ^a	P	log L	AIC	BIC	LRT ^b
MY					
1) LEG3,3_1	13	-3185.27	6396.54	6484.98	(2-1) 95.76**
2) LEG3,3_3	15	-3097.33	6224.66	6326.71	(3-2) 7.83**
3) LEG3,3_4	16	-3089.50	6211.01	6319.86	(4-3) 6.95 ^{ns}
4) LEG3,3_10	22	-3082.55	6209.11	6358.78	-
FY					
1) LEG3,3_1	13	1831.22	-3636.44	-3572.67	(2-1) 6.74**
2) LEG3,3_4	16	1837.95	-3643.89	-3565.40	(3-2) 1.55 ^{ns}
3) LEG3,3_6	18	1839.50	-3642.99	-3554.69	(4-3) 2.14 ^{ns}
4) LEG3,3_10	22	1841.64	-3639.27	-3531.34	-
PY					
1) LEG3,3_1	13	2389.90	-4753.80	-4690.18	(2-1) 16.97**
2) LEG3,3_4	16	2406.87	-4781.73	-4703.44	(3-2) 7.14 ^{ns}
3) LEG3,3_10	22	2414.00	-4784.01	-4676.35	-
MP					
1) LEG3,3_1	13	762.05	-1498.10	-1434.44	(2-1) 15.06**
2) LEG3,3_4	16	777.11	-1522.22	-1443.87	(3-2) 2.72 ^{ns}
3) LEG3,3_10	22	779.83	-1515.65	-1407.93	-

^a LEG_{k_s},k_{ap}-R or LEG_{k_s},k_{ap}-r: functions of Legendre polynomials (LEG) where k_s and k_{ap} correspond to the order of the covariance function for additive genetic and permanent environmental effects, respectively, and R and r indicate the structure of residual variance (homogeneous or heterogeneous, respectively).

^b Likelihood ratio test between hierarchical models. **p<0.01; ^{ns}: not significant.

The first three eigenvalues explained most of the additive genetic and permanent environmental variability (over 90%). On the basis of the results of the eigenvalues, the dimension of the three random effects can be reduced without the loss of information, disagreeing with the log L and AIC criteria for additive genetic and permanent environmental effects and agreeing with the BIC criterion for the additive genetic effect.

The estimates of additive genetic, permanent environmental, residual and phenotypic variances for test-day MY (LEG3,5_4), FY (LEG3,4_4), PY (LEG3,4_4), and MP (LEG3,4_4) are shown in figure 1. In general, the additive variances for the traits were higher in the first month, decreased until the eighth month, and then increased in the ninth and tenth month.

The estimates of permanent environmental variance for MY were practically constant across lactation and were higher than the additive variances. Residual variance was declined from the fifth to the tenth month (Figure 1). For PY, the additive genetic variances showed the same trend

Table 3. Number of estimated parameters (P), logarithm of the maximum likelihood function (log L), Akaike Information Criterion (AIC), Schwarz Bayesian information criterion (BIC) and Likelihood Ratio Test (LRT) for models used comparing the number of classes for the residual variance of Milk (MY), fat (FY), protein (PY) yields and mozzarella production (MP) in buffaloes.

Model ^a	P	log L	AIC	BIC	LRT ^b
MY					
1) LEG3,4_4	20	-3051.38	6142.77	6278.83	(2-1) 28.73**
2) LEG3,5_4	25	-3022.66	6095.32	6265.40 €	(3-2) 24.35**
3) LEG3,6_4	31	-2998.31	6058.62 €	6269.52	-
4) LEG4,4_4	24	-3049.33	6146.66	6309.94	(5-4) 28.25**
5) LEG4,5_4	29	-3021.08	6100.16	6297.45	(6-5) 24.92**
6) LEG4,6_4	35	-2996.16	6062.32	6300.26	-
7) LEG5,5_4	34	-3013.89	6095.78	6327.09	(8-7) 24.24**
8) LEG5,6_4	40	-2989.65	6059.30	6331.43	-
FY					
1) LEG3,4_4	20	1819.91	-3639.82 €	-3541.7 €	(2-1) 3.06 ^{ns}
2) LEG3,5_4	25	1811.85	-3623.69	-3501.05	(3-2) 7.99 ^{ns}
3) LEG3,6_4	31	1813.85	-3627.69	-3475.61	-
4) LEG4,4_4	24	1803.98	-3607.95	-3490.21	(5-4) 14.54**
5) LEG4,5_4	29	1813.51	-3627.03	-3484.76	(6-5) 2.9 ^{ns}
6) LEG4,6_4	35	1810.45	-3620.89	-3449.19	-
7) LEG5,5_4	34	1807.38	-3614.76	-3447.96	(8-7) 3.14 ^{ns}
8) LEG5,6_4	40	1804.52	-3609.04	-3412.81	-
PY					
1) LEG3,4_4	20	2386.79	-4733.58	-4635.71 €	(2-1) 9.55 ^{ns}
2) LEG3,5_4	25	2396.34	-4742.68	-4620.34	(3-2) 12.59 ^{ns}
3) LEG3,6_4	31	2408.93	-4755.87 €	-4604.13	-
4) LEG4,4_4	24	2392.47	-4736.94	-4619.50	(5-4) 2.21 ^{ns}
5) LEG4,5_4	29	2394.68	-4731.37	-4589.45	(6-5) 1.64 ^{ns}
6) LEG4,6_4	35	2396.32	-4742.65	-4591.37	-
7) LEG5,5_4	34	2386.2	-4704.39	-4538.01	(8-7) 10.43 ^{ns}
8) LEG5,6_4	40	2396.63	-4713.25	-4517.50	-
MP					
1) LEG3,4_4	20	777.36	-1514.72 €	-1416.78 €	(2-1) 0.97 ^{ns}
2) LEG3,5_4	25	778.32	-1506.64	-1384.23	(3-2) 5.00 ^{ns}
3) LEG3,6_4	31	783.29	-1504.57	-1352.77	-
4) LEG4,4_4	24	776.85	-1505.70	-1388.18	(5-4) 2.86 ^{ns}
5) LEG4,5_4	29	779.71	-1501.41	-1359.41	(6-5) 5.58 ^{ns}
6) LEG4,6_4	35	785.29	-1500.57	-1329.19	-
7) LEG5,5_4	34	779.95	-1491.90	-1325.41	(8-7) 4.07 ^{ns}
8) LEG5,6_4	40	784.02	-1488.04	-1292.17	-

€ Indicates the best model according to the AIC and BIC.

^a LEG_{k_s},k_{ap}-R or LEG_{k_s},k_{ap}-r: functions of Legendre polynomials (LEG) where k_s and k_{ap} correspond to the order of the covariance function for additive genetic and permanent environmental effects, respectively, and R and r indicate the structure of residual variance (homogeneous or heterogeneous, respectively).

^b Likelihood ratio test between hierarchical models. **p<0.01; ^{ns}: not significant.

Los tres primeros valores propios explican la mayor parte de la variabilidad aditiva genética y de ambiente permanente (más del 90%). En base a los resultados de los valores propios, es posible reducir la dimensión de los tres efectos aleatorios sin sufrir ninguna pérdida de información, en desacuerdo con los criterios de log L y AIC para los efectos aditivos genéticos y de

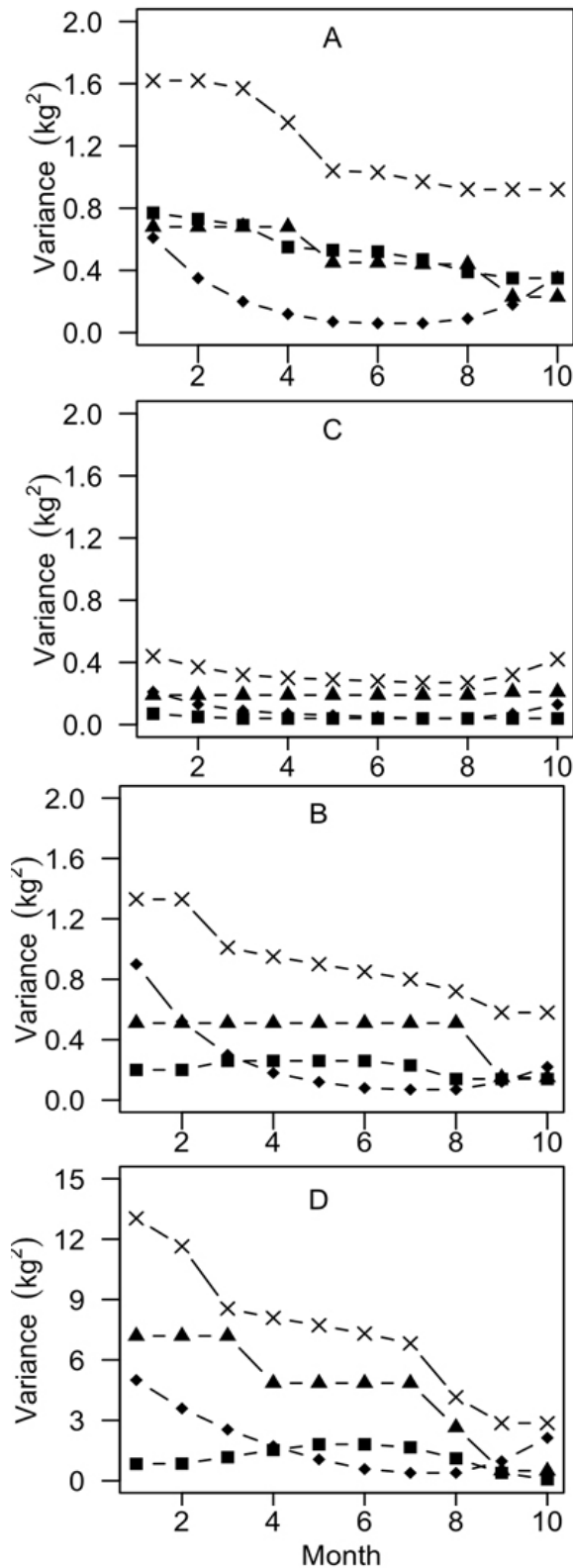


Figure 1. Estimates of additive (♦), permanent environmental (■), residual (▲) and phenotypic (x) variance for milk yield LEG3,5_4 (A), fat yield LEG3,4_4 (B), protein yield LEG3,4_4 (C), and mozzarella production LEG3,4_4 (D).

Table 4. Percentage of variance explained by the eigenvalues of the matrix of additive genetic and permanent environmental random regression coefficients for the models selected.

Traits	Random regression coefficient				
	1	2	3	4	5
MY LEG3,5_4					
Additive	84.92	15.06	0.02	-	-
Permanent environmental	69.94	19.93	7.12	2.35	0.67
FY LEG3,4_4					
Additive	79.36	15.48	5.16	-	-
Permanent environmental	82.10	12.66	2.92	2.32	-
PY LEG3,4_4					
Additive	62.65	22.60	14.75	-	-
Permanent environmental	62.87	16.25	11.14	9.74	-
MP LEG3,4_4					
Additive	88.90	10.43	0.67	-	-
Permanent environmental	97.52	1.79	0.45	0.24	-

ambiente permanente y de acuerdo con el criterio BIC para el efecto genético aditivo.

En la figura 1 se muestran las estimaciones de las varianzas aditivas genéticas, de ambiente permanente, residuales y fenotípicas para el día de control para MY (LEG3,5_4), FY (LEG3,4_4), PY (LEG3,4_4) y MP (LEG3,4_4). En general, las varianzas aditivas de los rasgos fueron mayores en el primer mes y disminuyeron hasta el octavo mes para luego aumentar en el noveno y décimo mes.

Las estimaciones de la varianza de ambiente permanente para MY fueron prácticamente constantes a lo largo del periodo de lactancia y fueron más altas que las varianzas aditivas. La varianza residual disminuyó desde el quinto al décimo mes (Figura 1).

Para PY, las varianzas genéticas aditivas mostraron la misma tendencia que las varianzas fenotípicas, observándose valores más altos al final de la lactancia. Para FY y MP, las varianzas genéticas aditivas mostraron la misma tendencia que las varianzas fenotípicas y residuales y fueron mayores al final del periodo de lactancia. Estas estimaciones más altas para FY, PY y MP pueden ser causadas por número de registros encontrados o debido al alto contenido de grasa y proteína al final de la lactancia (Figura 1).

Las heredabilidades para MY, FY, PY y MP variaron de 0.38 a 0.05, de 0.67 a 0.11, de 0.50 a 0.07 y de 0.50 a 0.08, respectivamente (Figura 2).

En la figura 3 se muestran las estimaciones de las correlaciones genéticas y fenotípicas en el día de control para MY, FY, PY y MP. Las correlaciones genéticas para estos rasgos variaron de -0,56 a 0,96. Las correlaciones fenotípicas de los rasgos estudiados fueron menores a medida que se incrementó el

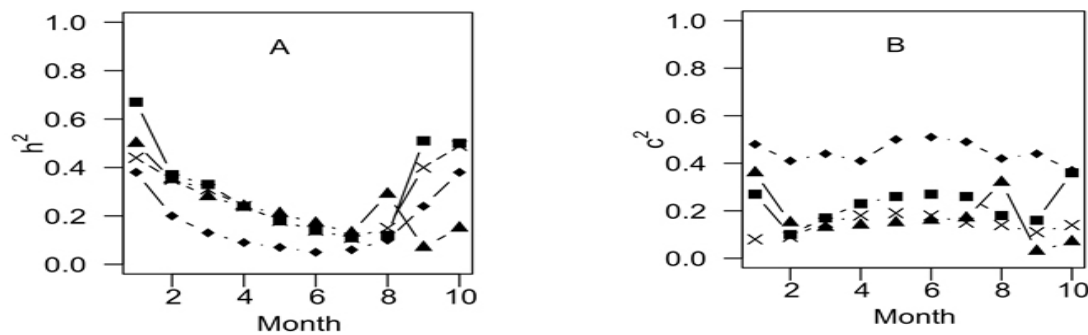


Figure 2. Estimates of heritability (h^2 , **A**) and proportion of phenotypic variance due to the permanent environment (c^2 , **B**) for milk yield LEG3,5_4 (♦), fat yield LEG3,4_4 (■), protein yield LEG3,4_4 (▲), and mozzarella production LEG3,4_4 (x).

as the phenotypic variances, with the observation of higher values at the end of lactation. For FY and MP, the additive genetic variances showed the same trend as the phenotypic and residual variances and were higher at the end of lactation. These higher estimates for FY, PY and MP might be due to the number of records found or to the high fat and protein content at the end of lactation (Figure 1).

The heritabilities for MY, FY, PY and MP ranged from 0.38 to 0.05, 0.67 to 0.11, 0.50 to 0.07 and 0.50 to 0.08, respectively (Figure 2).

The estimates of genetic and phenotypic correlations for test-day MY, FY, PY and MP are shown in figure 3. The genetic correlations for these traits ranged from -0.56 to 0.96. The phenotypic correlations of the traits studied were lower as the interval between test days increased (Figure 3). In general, the phenotypic

interval entre los días de control (Figura 3). En general, las correlaciones fenotípicas tuvieron una magnitud menor que las correlaciones genéticas.

DISCUSIÓN

La evaluación del mejor ajuste para la varianza residual (Tabla 2) mostró aumentos en log L ($p < 0.01$), que fueron significativos por el LRT, con un aumento en el número de clases heterogéneas. De acuerdo con los criterios utilizados para la evaluación de la bondad de ajuste (AIC y BIC), se consideró que el modelo que incorpora la homogeneidad de las varianzas residuales era inapropiado. Este hecho indica un comportamiento distinto de las varianzas residuales a lo largo del periodo de lactancia y la necesidad de considerar estructuras de varianza heterogénea para los efectos residuales. La heterogeneidad de las varianzas residuales puede atribuirse a factores como la etapa de preñez, la

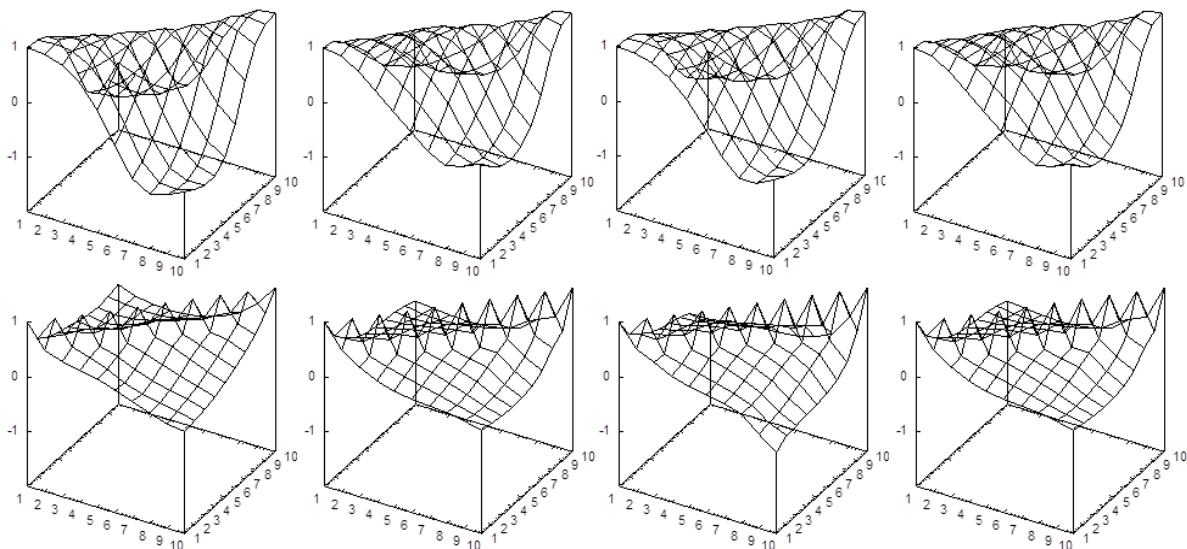


Figure 3. Estimates of genetic (top row) and phenotypic (bottom row) correlations for test-day milk yield (first column), protein yield (second column), fat yield (third column), and mozzarella production (fourth column) obtained by single-trait random regression.

correlations were of lower magnitude than the genetic correlations.

DISCUSSION

Evaluation of the best fit for residual variance (Table 2) showed increases of log L ($p < 0.01$), which were significant by the LRT, with increasing number of heterogeneous classes. According to the criteria used for the assessment of goodness-of-fit (AIC and BIC), the model considering homogeneity of residual variances was found to be inappropriate. This fact indicates a distinct behavior of residual variances across lactation and the need to consider heterogeneous variance structures for residuals effects. The heterogeneity of residual variances can be attributed to factors such as stage of pregnancy, body condition, duration of the lactation interval and management conditions, since these factors are not easily incorporated into models analyses because of the lack of information (6-7).

The changes in log L were of small magnitude and nonsignificant by the LRT ($p > 0.01$), indicating that four residual classes would be sufficient to obtain a good fit of residual variance for MY, PY, FY, and MP. This would avoid the use of superparameterized models which generally present parameter estimation problems (7).

The BIC, which is more rigorous because of parameterization, indicated that LEG3,5_4, LEG3,4_4, LEG3,4_4 and LEG3,4_4 were the best fitted models for MY, FY, PY and MP, respectively. Therefore, these models would be more adequate to describe the biological variation in yield traits across lactation.

Specifically for PY and FY, the increases seen in late lactation may be related mainly to the small number of records available for this period (Table 1) or may have been due to computing artifacts rather than biology (8). In addition, fitting RRM at the beginning and end of lactation is difficult, resulting in parameter estimation problems (3-4). The results might be explained by the fact that buffalo breeding programs are only starting in Colombia and the population therefore presents small selection rates for the traits studied. In addition, the low yields in early and late lactation are related to the lack of feeding and management strategies during these periods.

The trends of additive and permanent environmental variances across lactation obtained for the traits studied (MY, PY and FY) are comparable to those reported by Jamrozik and Schaeffer (4) and Silvestre et al (5), who

condición corporal, la duración del intervalo de lactancia y las condiciones de manejo, ya que estos factores no pueden incorporarse fácilmente en el análisis de los modelos debido a la falta de información (6-7).

Los cambios en log L fueron de una magnitud pequeña y no significativa por medio de LRT ($p > 0.01$), lo que indica que sería suficiente utilizar cuatro clases residuales para obtener un buen ajuste de la varianza residual para MY, PY, FY y MP. Esto evitaría el uso de modelos superparametrizados que generalmente presentan problemas en la estimación de los parámetros (7).

El BIC, el cual es más riguroso debido a la parametrización, indicó que LEG3,5_4, LEG3,4_4, LEG3,4_4 y LEG3,4_4 fueron los modelos con el mejor ajuste para MY, FY, PY y MP, respectivamente. Por lo tanto, estos modelos serían más adecuados para describir la variación biológica de las características productivas a lo largo de la lactancia.

Específicamente para PY y FY, los aumentos al final de la lactancia pueden estar relacionados principalmente con un menor número de registros disponibles para este período (Tabla 1) o pueden ser causado por problemas computacionales en lugar de la biología (8). Además, resulta difícil realizar el ajuste del MRA al inicio y al final de la lactancia, lo que ocasiona problemas en la estimación de los parámetros (3-4). Los resultados pueden explicarse por el hecho de que los programas de cría de búfalos hasta ahora están empezando en Colombia y por lo tanto, la población posee tasas pequeñas de selección de características estudiadas. Adicionalmente, la baja producción al inicio y al final de la lactancia se relacionan con la falta de estrategias de alimentación y manejo durante estos períodos.

Las tendencias de las varianzas aditivas y de ambiente permanente a lo largo de la lactancia obtenidas para MY, PY y FY son comparables con aquellas reportadas por Jamrozik y Schaeffer (4) y Silvestre et al (5), quienes estimaron varianzas aditivas y de ambiente permanente más altas al principio y al final de la lactancia para MY, PY y FY en bovinos lecheros.

Las heredabilidades para MY oscilaron entre 0,38 (primer mes) y 0.05 (séptimo mes). Sesana et al (7) encontró estimaciones más altas, particularmente en los extremos de la curva de lactancia, un hallazgo que puede atribuirse al hecho de que los autores utilizaron registros semanales. Las heredabilidades para FY variaron entre 0.67 (primer mes) y 0.11 (séptimo

estimated higher additive and permanent environmental variances at the beginning and end of lactation for MY, PY and FY in dairy cows.

The heritabilities for MY ranged from 0.38 (first month) to 0.05 (seventh month). Sesana et al (7) found higher estimates, particularly at the extremes of the lactation curve, a finding that might be attributed to the fact that the authors used weekly records. The heritabilities for FY ranged from 0.67 (first month) to 0.11 (seventh month), whereas the heritabilities for PY ranged from 0.50 (first month) to 0.07 (ninth month). Similar results have been reported for dairy cattle using RRM, with high estimates for FY and PY at the beginning and end of lactation (4-6, 9-11). Aspilcueta et al (6) reported heritabilities of lower magnitude using finite dimension models for monthly records of FY and PY in dairy buffaloes. The heritabilities for MP ranged from 0.50 (first month) to 0.08 (ninth month). Rosati and Van Vleck (12) reported a heritability for MP of 0.13 for 270-day lactation in the Italian buffalo population. In general, the heritability estimates for MY, PY, FY and MP showed a similar trend across lactation, with the observation of a marked increase at the end of this period (Figure 2). Similar trends have been reported for dairy cattle (4-5).

For MY, the fractions of phenotypic variance due to permanent environmental effects were higher than the heritability estimates across lactation. The fractions of phenotypic variance for FY and PY were higher in the first month and between the fourth and seventh month of lactation, respectively. For MP, the fraction of phenotypic variance due to permanent environmental effects remained constant across lactation.

The estimates were higher between adjacent test days (early lactation) and decreased with increasing distance between test days, especially between tests performed at the beginning and end of lactation. Negative and near to zero genetic correlations between milk tests performed at the beginning and end of lactation have been reported for dairy cattle (13-14). Sesana et al (7) and Breda et al (15) reported similar trends in genetic correlations for buffaloes.

The correlations, variances and heritabilities obtained for the traits studied are probably due to the difficulty in modeling yields in early lactation, a phase during which the buffalo cow suffers from postpartum stress, as well as negative energy balance and nutritional deficiencies. In addition, genetic breeding programs of dairy buffaloes are only starting in Colombia and the herds and environment where the animals are

mes), mientras que las heredabilidades para PY oscilaron entre 0.50 (primer mes) y 0.07 (noveno mes). Se han reportado resultados similares para bovinos lecheros usando MRA con estimaciones altas para FY y PY al principio y al final de la lactancia (4-6, 9-11). Aspilcueta et al (6) reportó las heredabilidades de una menor magnitud utilizando modelos de dimensión finita para registros mensuales de FY y PY en búfalos lecheros. Las heredabilidades para MP variaron de 0.50 (primer mes) a 0.08 (noveno mes). Rosati y Van Vleck (12) reportaron una heredabilidad para MP de 0.13 para un periodo de lactancia de 270 días en la población de búfalos de Italia. En general, las estimaciones de la heredabilidad para MY, PY, FY y MP mostraron una tendencia similar durante la lactancia, observando un aumento marcado al final de este periodo (Figura 2). Se han reportado tendencias similares para bovinos lecheros (4-5).

Para MY, las fracciones de la varianza fenotípica debido a efectos de ambiente permanente fueron más altas que las estimaciones de heredabilidad a lo largo de la lactancia. Las fracciones de la varianza fenotípica para FY y PY fueron más altas en el primer mes y entre el cuarto y séptimo mes de lactancia, respectivamente. Para MP, la fracción de la varianza fenotípica debido a efectos de ambiente permanente se mantuvo constante durante la lactancia.

Las estimaciones fueron mayores entre días de control consecutivos (inicio lactancia) y disminuyeron al aumentar la distancia entre los días de control, especialmente entre los controles realizados al principio y al final de la lactancia. Se han reportado correlaciones genéticas negativas y cercanas a cero entre el control de leche realizado al principio y al final de la lactancia para bovinos lecheros (13-14). Sesana et al (7) y Breda et al (15) informaron tendencias similares en correlaciones genéticas de búfalos.

Las correlaciones, varianzas y heredabilidades obtenidas para las características estudiadas son causadas posiblemente a la dificultad de modelar la producción al inicio de la lactancia, una fase durante la cual la búfala sufre de estrés postparto, así como un balance energético negativo y deficiencias nutricionales. Además, los programas de mejoramiento genético en búfalos lecheros, hasta ahora están iniciando en Colombia y el ambiente donde se crían los animales y los rebaños, así como las bajas tasas de selección y un número pequeño de observaciones, puede influir en las etapas tempranas y tardías de la lactancia. Esto sugiere el uso de períodos más cortos de lactancia para las poblaciones de búfalos cuando los programas

raised, as well as the low selection rates and small number of observations may influence the early and late stages of lactation. This suggests the use of shorter lactation periods for buffalo populations in which selection programs are only started. De Roos et al (10) suggested the need to model fixed curves since the estimates of genetic variances at the beginning and end of lactation are increased if fixed herd curves are not modeled independently.

In conclusion, the RRM using LP are appropriate to describe the genetic variation in test-day MY, FY, PY and MP of buffaloes. Problems with the estimation of genetic parameters were observed at the beginning and at the end of the lactation curve, indicating that the traits studied may be affected by the number of available records at the extremes of the curve or by the high content of fat and protein, or may have been due to computing artifacts rather than biology (8). Residual variances varied across lactation and can be modeled by forming heterogeneous classes using low-order LP, explaining the changes in genetic variation according to days in milk.

Acknowledgments

This research had financial support from the Colombian Ministry of Agriculture and Rural Development, Colombian Cattle Breeders Federation (FEDEGAN), Colombian Buffaloes Breeders Association, Codi sostenibilidad E01808 and Fundação de Apoio à Pesquisa do Estado de São Paulo (FAPESP- Part of doctoral thesis of first author: Process N°2009/53773-1).

de selección estén empezando. De Roos et al (10) sugirieron la necesidad de modelar curvas fijas ya que las estimaciones de las varianzas genéticas al principio y al final del periodo de lactancia se incrementan, si las curvas fijas del rebaño no han sido modeladas de forma independiente.

En conclusión, la utilización de MRA via LP es apropiada para describir la variación genética en el día de control para MY, FY, PY y MP de búfalos. Se observaron problemas relacionados con la estimación de los parámetros genéticos al principio y al final de la curva de lactancia, lo que indica que las características estudiadas pueden ser afectadas por el número de registros disponibles en los extremos de la curva o por el alto contenido de grasa y proteína, o puede deberse a los problemas computacionales en lugar de la biología (8). Las varianzas residuales oscilaron durante la lactancia y pueden ser modeladas mediante clases heterogéneas usando LP de orden inferior, explicando los cambios en la variación genética de acuerdo con los días de lactancia.

Agradecimientos

Esta investigación contó con el apoyo financiero del Ministerio Colombiano de Agricultura y Desarrollo Rural, la Federación Colombiana de Ganaderos (FEDEGAN), la Asociación Colombiana de Criadores de Búfalos, Codi sostenibilidad E01808 y Fundação de Apoio à Pesquisa do Estado de São Paulo (FAPESP – Parte de la tesis doctoral del primer autor: Proceso No. 2009/53773-1).

REFERENCES

1. Tonhati H, Lima A, Lanna D, Camargo G, Baldi F, Albuquerque L, Montrezor J. Milk fatty acid characterization and genetic parameter estimates for milk conjugated linoleic acid in buffaloes. *J Dairy Res* 2011; 78(2):178–183.
2. Seno LO, Cardoso V, Tonhati H. Valores econômicos para as características de produção de leite de búfalas no Estado de São Paulo. *R Bras Zootec* 2007; 36(6):2016–2022.
3. El Faro L, Cardoso Vera L, Albuquerque LG. Variance component estimates applying random regression models for test-day milk yield in Caracu heifers (*Bos taurus Artiodactyla*, *Bovidae*). *Genet Mol Biol* 2008; 31(3):665-673.
4. Jamrozik J, Schaeffer LR. Estimates of Genetic Parameters for a Test Day Model with Random Regressions for Yield Traits of First Lactation Holsteins. *J Dairy Sci* 1997; 80(4):762–770.
5. Silvestre A, Petim-Batista F, Colaço J. Genetic parameter estimates of Portuguese dairy cows for milk, fat and protein using a spline test-day model. *J Dairy Sci* 2005; 88(3):1225-1230.
6. Aspilcueta RB, Sesana RC, Muñoz-Berrocal M, Seno LO, Bignardi AB, El Faro L, et al. Genetic parameters for milk, fat and protein yields in Murrah buffaloes (*Bubalus bubalis Artiodactyla*, *Bovidae*). *Genet Mol Biol* 2010; 33(1):71-77.

7. Sesana RC, Bignardi AB, Aspilcueta RA, El Faro L, Baldi F, Albuquerque LG, Tonhati H. Random regression models to estimate genetic parameters for test-day milk yield in Brazilian Murrah buffaloes. *J Anim Breed Genet* 2010; 127(5):369–376.
8. Martínez CA, Elzo M, Manrique C, Grajales LF, Jimenez A. Random Regression Models for Estimation of Covariance Functions, Genetic Parameters and Prediction of Breeding Values for Rib Eye Area in a Colombian *Bos indicus*-*Bos taurus* Multibreed Cattle Population. *Rev Colomb Estad* 2012; 35(2):309-330.
9. Borghese A, Mazzi M. Buffalo Production And Research. Rome: Editorial Food and Agriculture Organization of The United Nations; 2005.
10. Meyer KWOMBAT. A tool for mixed model analyses in quantitative genetics by REML, J Zhejiang Uni 2007; 8:815-821.
11. De Roos APW, Harbers AGF, de Jong G. Random herd curves in a test-day model for milk, fat, and protein production of dairy cattle in The Netherlands. *J Dairy Sci* 2004; 87(8): 2693–2701.
12. Rosati A, Van Vleck LD. Estimation of genetic parameters for milk, fat, protein and mozzarella cheese production for the Italian river buffalo *Bubalus bubalis* population. *Livest Prod Sci* 2002; 74 (2):203–217.
13. Munera BOD, Herrera RAC, Cerón-Muñoz MF. Componentes de (co) varianza y parámetros genéticos para producción y valor económico de la leche a través de modelos de regresión aleatoria en hembras Holstein de primera lactancia. *Livest Res Rural Dev* 2013;25(12): art 208.
14. Herrera RAC, Munera OD, Cerón-Muñoz MF. Variance components and genetic parameters for milk production of Holstein cattle in Antioquia (Colombia) using random regression models. *Rev Col Cien Pec* 2013; 26:90-97.
15. Breda F, Albuquerque L, Eudydes RF, Bignardi A B, Baldi F, Torres RA, Barbosa L, Tonhati H. Estimation of genetic parameters for milk yield in Murrah buffaloes by bayesian inference. *J Dairy Sci* 2010; 93(2):784-791.