Genetic parameters for milk production by using random regression models with different alternatives of fixed regression modeling

Jaime Araújo Cobuci[1,3], Claudio Napolis Costa[2,3], José Braccini Neto[1], Ary Ferreira de Freitas[2,3]

[1] Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil.
[2] Embrapa Gado de Leite, Juiz de Fora, MG, Brasil.
[3] Bolsista do CNPq.

ABSTRACT - Records of test-day milk yields of the first three lactations of 25,500 Holstein cows were used to estimate genetic parameters for milk yield by using two alternatives of definition of fixed regression of the random regression models (RRM). Legendre polynomials of fourth and fifth orders were used to model regression of fixed curve (defined based on averages of the populations or multiple sub-populations formed by grouping animals which calved at the same age and in the same season of the year) or random lactation curves (additive genetic and permanent enviroment). Akaike information criterion (AIC) and Bayesian information criterion (BIC) indicated that the models which used multiple regression of fixed lactation curves of lactation multiple regression model with fixed lactation curves had the best fit for the first lactation test-day milk yields and the models which used a single regression of fixed curve had the best fit for the second and third lactations. Heritability for milk yield during lactation estimates did not vary among models but ranged from 0.22 to 0.34, from 0.11 to 0.21, and from 0.10 to 0.20, respectively, in the first three lactations. Similarly to heridability estimates of genetic correlations did not vary among models. The use of single or multiple fixed regressions for fixed lactation curves by RRM does not influence the estimates of genetic parameters for test-day milk yield across lactations.

Key Words: test-day milk yield, genetic correlation, heritability, Legendre polynomial, selection

Parâmetros genéticos para produção de leite usando modelos de regressão aleatória com diferentes alternativas de modelagem da regressão fixa

RESUMO - Os registros de produção de leite no dia do controle das três primeiras lactações de 25,5 mil vacas da raça Holandesa foram utilizados para estimar parâmetros genéticos para produção de leite usando duas alternativas de definição da regressão fixa dos modelos de regressão aleatória (MRA). Os polinômios de Legendre de ordens 4 e 5 foram usados para modelar as regressões das curvas fixas (definidas com base nas médias das produções de leite no dia do controle da população ou de múltiplas sub-populações formadas pelo agrupamento de animais que pariram na mesma idade e estação do ano) e aleatórias (genética aditiva e de ambiente permanente) de lactação. Os critérios de informação de Akaike (AIC) e Bayesiano (BIC) indicaram os modelos que consideraram múltiplas regressões das curvas fixas de lactação como os que melhor se ajustaram aos registros de produção de leite da primeira lactação e os modelos que utilizaram uma única regressão da curva fixa, como os melhores para ajuste das segunda e terceira lactações. As herdabilidades para produção de leite ao longo da lactação não variaram entre modelos, entretanto variaram de 0,22 a 0,34; 0,11 a 0,21 e 0,10 a 0,20, respectivamente, para as três primeiras lactações. Semelhantemente às estimativas de herdabilidade os valores das estimativas de correlações genéticas não variaram entre modelos. O uso de uma ou de múltiplas regressões das curvas fixas de lactação pelos MRA não influencia na estimação de parâmetros genéticos para as produções de leite ao longo das lactações de vacas da raça Holandesa.

Palavras-chave: correlação genética, herdabilidade, polinômio de Legendre, produção de leite no dia do controle, seleção

Introduction

The use of more precise model definitions in genetic evaluations contributes to increase efficiency of selection programs. Improvements in modeling of environmental effects that influence the productive traits of dairy cattle can be achieved by using the test-day models. These models fundamentally evaluate differences in the lactation curves of animals (Bormann et al., 2003). Among the models that consider test-day production, random regression model has been widely shown to increase the accuracy of breeding value predictions and has been already implemented in the

official genetic evaluations of dairy cows in many countries (Strabel et al. 2004).

The use of random regression models with parametric functions as covariates, allows to split the shape of lactation curve into two parts: a general part (fixed) for assessing similarities of the lactation curves within specific groups of animals (i.e.: similar age, stage of lactation, parity and season of birth) and a second part specific for each individual animal (random) (Bormann et al., 2003).

Choosing the most appropriate model depends on decisions concerning the effects that are to be included in the model, especially those accounting for similarities of lactation curves of groups of animals under the influence of a common environmental effect (fixed part) or individual lactation curves of animals (random part). In the fixed part, one or multiple regressions of lactation fixed curves have been evaluated (Strabel et al., 2004, Costa et al., 2008, Muir et al., 2007).

The present study aimed to one or multiple regressions of fixed lactation curves included in random regression models used for the estimation of genetic parameters of test-day milk for the first three lactations of Holstein cows in Brazil.

Material and Methods

Data consisted of 2.03 million test-day records made available by the Associação Brasileira de Criadores de Gado Holandês (ABCBRH). Only records of the first three lactations of cows calving between 1993 and 2004 were considered in the analyses. Test day prior to the 5th day and after the 305th day of lactation were excluded. Age range by parity was 20 to 48 months for first lactation, 33 to 67 months for second lactation and 45 to 87 months for third lactation. Cows were required to have a minimum of six test-day records per lactation. Test-day observations were deleted if a herd-year-month test day contained less than four observations. After these restrictions, 25,528, 11,767 and 4,265 records, respectively, from the first, the second and the third lactation of 41,560 cows remained available for analyses with random regression models using Legendre polynomials of orders four and five (Table 1).

Four classes of age at calving (20 to 24, 25 to 29, 30 to 34 and 35 to 48 months) and four calving seasons (January through March, April through June, July through September and October through December) were combined to produce 16 age-season classes. Two models were fitted to the data. In the first model it was considered the average test-day milk yield of each sub-population which calved within the same age-season class, defining the regression of fixed lactation curves (multiple regressions) as an alternative to the model that considered the average test-day milk yield of the population defining the regression of fixed lactation curve (one single regression). The models were designated by U4, M4, M5 or U5 when referring to Legendre polynomials of orders four or five used in one (U) or multiple (M) regressions of fixed lactation curves analyses.

The random regression equations U, assuming the same sub model to fit fixed (average population test-day milk yield lactation curve), genetic and permanent environment effects, was

$$y_{ijkl} = RAMC_i + E_j + \sum_{m=1}^{1} b_m x_{ijkl} + \sum_{m=1}^{n} q_{km} Z_{klm} + \sum_{m=1}^{n} a_{km} Z_{klm} +$$

$$+ \sum_{m=1}^{n} p_{km} Z_{klm} + e_{ijkl} \qquad (U),$$

where y_{ijk1} = milk yield of cow k in the lactation 1, in period of lactation t within the classes i (herd-year-month test-day) and j (season of calving); $RAMC_i$ = fixed effect of herd-year-month test-day; E_j = calving season j; b_m = linear regression coefficient of milk yield on age at calving; x_{ijkl} = age at calving; q_{km} = vector of fixed regression coefficients specific to modeling the average lactation curve of the population; a_{km} and p_{km} = vectors of random regression coefficients that describe, respectively, the additive genetic and permanent environmental effects; Z_{klm} = n-th parameter of Legendre polynomials of order 4 or 5; e_{ijkl} = random residual effect associated to y_{ijkl}.

Similarly, the random regression equations M, assuming the same sub model to fit fixed (average age-season test-day milk yield lactation curves), genetic and permanent environment effects, was:

Table 1 - Description of the data set used in the study

Parameter	Lactation 1	Lactation 2	Lactation 3
Average daily milk (standard deviation, kg)	23.08 (6,77)	26.94 (8.50)	28.58 (8.93)
Average age of calving (standard deviation, months)	28.4 (4,6)	42.9 (5.9)	57.3 (7.1)
Number of lactations	25,528	11,767	4,265
Number of test-day records	225,357	101,913	36,624
Average number of test-day per lactations	8.8	8.6	8.6
Number of herds	384	248	130
Number of records per contemporary group	13.32	7.17	3.21
Number of animals in the relationship matrix	42,347	20,845	8,106
Average number of daughters per bull	10.59	9.7	8.9

Genetic parameters for milk production by using random regression models with ...

559

$$y_{ijkl} = RAMC_i + \sum_{m=1}^{n} \beta_{jm} Z_{klm} + \sum_{m=1}^{n} a_{km} Z_{klm} +$$

$$+ \sum_{m=1}^{n} p_{km} Z_{klm} + e_{ijkl} \tag{M}$$

where y_{ijk1} = milk yield of cow k in the lactation 1, in period of lactation t within the classes i (herd-year-month test-day) and j (age-season of calving); $RAMC_i$ = fixed effect of herd-year-month test-day; β_{jm} = vector of fixed regression coefficients specific to modeling the average lactation curve of age-season of calving classes; a_{km} and p_{km} = vectors of random regression coefficients that describe, respectively, the additive genetic and permanent environmental effects; e_{ijkl} = residual random effect associated with y_{ijkl}; Z_{jlm} represents the nth parameter Legendre polynomial of order 4 or 5, used in the description of genetic and permanent environment, as well as, in modeling the fixed lactation curves (averages) of animals belonging to classes of age-season

of calving. The temporary environmental (residual) variance was assumed homogeneous throughout the period of lactation.

The variance and covariance components of regression coefficients for random additive genetic and permanent environmental effects were estimated using the REMLF90 program (Misztal, 2005). Convergence was assumed when the difference between the -2log values of the likelihood functions obtained in consecutive iterations was smaller than 10^{-11}. Fit of different models was compared by examining the value of the maximum likelihood function (-2LogL), the Akaike information criterion-AIC (Akaike, 1973) and the Bayesian information criterion-BIC (Schwarz, 1978).

Results and Discussion

Covariance estimates between random regression additive and permanent environment coefficients estimated

Table 2 - Estimates of covariance components for random regression coefficients obtained by single regression models

	Model U4			Model U5		
	Lactation 1	Lactation 2	Lactation 3	Lactation 1	Lactation 2	Lactation3
Additive genetic regression coefficients						
a_1a_1	11.58	10.29	11.40	11.71	10.22	11.48
a_1a_2	1.569	1.352	0.7022	1.675	1.381	0.9148
a_1a_3	-1.318	-0.6493	-1.207	-1.368	-0.7087	-1.218
a_1a_4	0.1078	-0.2708	-0.8712	0.08761	-0.1913	-0.9005
a_1a_5	-	-	-	-0.1158	-0.4030	-0.3350
a_2a_2	2.180	1.786	0.8849	2.167	1.925	0.9247
a_2a_3	-0.8450	-0.09784	0.1938	-0.8418	-0.02268	0.1608
a_2a_4	0.002063	-0.1949	-0.09241	0.003514	-0.1502	-0.02614
a_2a_5	-	-	-	0.01210	0.01522	-0.1576
a_3a_3	0.7235	0.4711	0.8579	0.7475	0.4851	0.8455
a_3a_4	-0.06788	-0.1320	0.05036	-0.07102	-0.1384	0.1128
a_3a_5	-	-	-	0.04022	0.03455	-0.05681
a_4a_4	0.1506	0.2189	0.3304	0.1319	0.2206	0.3940
a_4a_5	-	-	-	-0.04682	-0.03932	-0.02898
a_5a_5	-	-	-	0.06699	0.07986	0.09596
Permanent environment regression coefficients						
p_1p_1	18.90	36.28	41.65	18.83	36.49	41.79
p_1p_2	-0.4495	-1.428	-1.156	-0.4760	-1.370	-1.283
p_1p_3	-0.8499	-2.536	-2.582	-0.9430	-2.528	-2.621
p_1p_4	0.09861	-0.2597	0.4626	0.1661	-0.1188	0.6723
p_1p_5	-	-	-	-0.5530	-0.4301	-0.7332
p_2p_2	4.224	7.486	8.807	4.344	7.552	8.983
p_2p_3	0.1387	-0.2034	-0.8031	0.1866	-0.1939	-0.6970
p_2p_4	-0.3367	-0.3451	-0.6069	-0.3441	-0.3574	-0.6277
p_2p_5	-	-	-	-0.1088	-0.4056	-0.1632
p_3p_3	1.912	3.550	3.991	1.835	3.436	3.875
p_3p_4	-0.1322	-0.1869	-0.5411	-0.1155	-0.1649	-0.5954
p_3p_5	-	-	-	-0.2674	-0.4319	-0.4973
p_4p_4	0.9980	1.466	2.045	0.9901	1.414	1.933
p_4p_5	-	-	-	-0.1356	-0.2683	-0.3323
p_5p_5	-	-	-	0.5175	0.7917	0.9096

U4 and U5 - models using Legendre polynomials of orders four and five and including single (U) regression of fixed lactation curve.

Table 3 - Estimates of covariance components of the random regression coefficients obtained by multiple regressions models

	Model M4			Model M5		
	Lactation 1	Lactation 2	Lactation 3	Lactation 1	Lactation 2	Lactation 3
Additive genetic regression coefficients						
a_1a_1	11.51	10.24	10.70	11.61	10.11	10.73
a_1a_2	1.270	1.238	0.7514	1.373	1.321	0.9621
a_1a_3	-1.248	-0.6450	-1.012	-1.289	-0.6761	-1.013
a_1a_4	0.1103	-0.1945	-0.6929	0.08563	-0.1507	-0.7470
a_1a_5	-	-	-	-0.1368	-0.4217	-0.3627
a_2a_2	1.879	1.801	0.8725	1.876	1.915	0.9486
a_2a_3	-0.7080	-0.05046	0.1217	-0.7003	-0.008831	0.1192
a_2a_4	-0.01542	-0.1959	-0.09917	0.006220	-0.1437	-0.01849
a_2a_5	-	-	-	0.02197	0.01563	-0.1509
a_3a_3	0.6737	0.4754	0.7830	0.6917	0.4822	0.7893
a_3a_4	-0.06123	-0.1533	0.005942	-0.06814	-0.1439	0.07075
a_3a_5	-	-	-	0.03447	0.03839	-0.04298
a_4a_4	0.1442	0.2567	0.3214	0.1365	0.2233	0.3921
a_4a_5	-	-	-	-0.04599	-0.04251	-0.03793
a_5a_5	-	-	-	0.06723	0.08220	0.09771
Permanent environment regression coefficients						
p_1p_1	18.94	36.28	42.19	18.89	36.54	42.38
p_1p_2	-0.2829	-1.372	-1.063	-0.3144	-1.346	-1.167
p_1p_3	-0.8781	-2.505	-2.704	-0.9781	-2.518	-2.736
p_1p_4	0.09182	-0.2645	0.3739	0.1629	-0.1048	0.5966
p_1p_5	-	-	-	-0.5351	-0.4057	-0.6589
p_2p_2	4.323	7.466	8.678	4.433	7.542	8.819
p_2p_3	0.07177	-0.2370	-0.7630	0.1186	-0.1989	-0.6892
p_2p_4	-0.3338	-0.3455	-0.6667	-0.3426	-0.3674	-0.7034
p_2p_5	-	-	-	-0.1046	-0.4030	-0.1700
p_3p_3	1.937	3.544	4.077	1.864	3.438	3.945
p_3p_4	-0.1333	-0.1738	-0.5070	-0.1142	-0.1649	-0.5664
p_3p_5	-	-	-	-0.2663	-0.4278	-0.5095
p_4p_4	0.9945	1.419	1.995	0.9796	1.393	1.873
p_4p_5	-	-	-	-0.1331	-0.2582	-0.3302
p_5p_5	-	-	-	0.5142	0.7984	0.9107

M4 and M5 - models using Legendre polynomials of orders four and five and including multiple (M) regressions of fixed age-season of calving lactation curves.

by models M4 and M5 were lower than those obtained by U4 and U5 models (Tables 2 and 3). For those models, regardless of the lactation, either for additive genetic or for permanent environment effects, the highest variance estimates were associated with the first coefficient (a_1 and p_1).

The correlations between regression coefficients for both additive genetic and permanent environment effects ranged from -0.67 to 0.33 for the first, from -0.46 to 0.31 for the second and from -0.53 to 0.30 for the third lactation (Tables 4 and 5). The largest values were observed for the additive genetic effect. In general, correlation estimates were similar among models U4, U5, M4 and M5, indicating no difference in the estimates of correlations between the random regression coefficients when modeling single or multiple regression fixed lactation curves.

Araújo et al. (2006), in Brazil, and Pool et al. (2000), in the Netherlands, reported correlations among regression coefficients estimated with Legendre polynomials of order 4 and 5, ranging from 0.57 to -0.54 and from -0.53 to 0.35,

respectively, for milk production of first lactation of Holstein cows.

The interesting mathematical properties of orthogonal polynomials (Strabel & Jamrozik, 2006) have supported this function to be considered as the most suitable for fitting test-day milk yield data (Schaeffer, 2004). However, the biological interpretation about their coefficients is more difficult than other parametric functions, such as the Wilmink function and the polynomial function of Ali & Schaeffer. Strabel & Jamrozik (2006) reported that the first Legendre polynomial coefficient (a_1), associated with the additive genetic effect, is related to total milk production and the second (a_2) is related to persistence in the lactation curve. Accordingly to this interpretation, the genetic correlation estimates between the additive genetic regression coefficients (a_1 and a_2) indicate low genetic association between total milk production and persistency. The estimates from M4 and M5 models were lower than those obtained from the U4 and U5 models (Tables 4 and 5).

Table 4 - Estimates of correlations between the random regression coefficients obtained by single regression models

	Model U4			Model U5		
	Lactation 1	Lactation 2	Lactation 3	Lactation 1	Lactation 2	Lactation3
Additive genetic regression coefficients						
a_1a_2	0.32	0.31	0.22	0.33	0.31	0.28
a_1a_3	-0.45	-0.29	-0.38	-0.46	0.32	-0.39
a_1a_4	0.08	-0.18	-0.44	0.07	0.12	-0.42
a_1a_5	-	-	-	-0.13	-0.44	-0.32
a_2a_3	-0.67	-0.10	0.22	-0.66	-0.02	0.18
a_2a_4	-0.00	-0.31	-0.17	0.01	-0.23	-0.04
a_2a_5	-	-	-	0.03	0.04	-0.53
a_3a_4	-0.20	-0.41	0.09	-0.22	-0.42	0.19
a_3a_5	-	-	-	0.18	0.17	-0.20
a_4a_5	-	-	-	-0.49	-0.29	-0.15
Permanent environment regression coefficients						
p_1p_2	-0.05	-0.09	-0.06	-0.05	-0.08	-0.06
p_1p_3	-0.14	-0.22	-0.20	-0.16	-0.22	-0.20
p_1p_4	0.02	-0.03	0.05	0.04	-0.02	0.07
p_1p_5	-	-	-	-0.17	-0.08	-0.12
p_2p_3	0.05	-0.04	-0.13	0.06	-0.03	-0.12
p_2p_4	-0.16	-0.10	-0.14	-0.16	-0.11	-0.15
p_2p_5	-	-	-	-0.07	-0.16	-0.06
p_3p_4	-0.09	-0.08	-0.19	-0.08	-0.07	-0.22
p_3p_5	-	-	-	-0.27	-0.26	-0.26
p_4p_5	-	-	-	-0.19	-0.25	-0.25

U4 and U5 - models using Legendre polynomials of orders four and five and including single (U) regression of fixed lactation curve.

Table 5 - Estimates of correlations between the random regression coefficients obtained by multiple regression models

	Model M4			Model M5		
	Lactation 1	Lactation 2	Lactation 3	Lactation 1	Lactation 2	Lactation 3
Additive genetic regression coefficients						
a_1a_2	0.27	0.29	0.25	0.29	0.30	0.30
a_1a_3	-0.45	-0.29	-0.35	-0.45	-0.31	-0.35
a_1a_4	0.09	-0.12	-0.37	0.07	-0.10	-0.36
a_1a_5	-	-	-	-0.15	-0.46	-0.35
a_2a_3	-0.63	-0.05	0.15	-0.61	-0.01	0.14
a_2a_4	-0.03	-0.29	-0.19	-0.01	-0.22	-0.03
a_2a_5	-	-	-	0.06	0.04	-0.50
a_3a_4	-0.20	-0.44	-0.01	-0.22	-0.44	0.12
a_3a_5	-	-	-	0.16	0.19	-0.15
a_4a_5	-	-	-	-0.48	-0.32	-0.19
Permanent environment regression coefficients						
p_1p_2	-0.03	-0.08	-0.06	-0.03	-0.08	-0.06
p_1p_3	-0.15	-0.22	-0.21	-0.16	-0.22	-0.21
p_1p_4	0.02	-0.04	0.04	-0.04	-0.01	0.07
p_1p_5	-	-	-	-0.17	-0.08	-0.11
p_2p_3	0.02	-0.05	-0.13	0.04	-0.04	-0.12
p_2p_4	-0.16	-0.11	-0.16	-0.16	-0.11	-0.17
p_2p_5	-	-	-	-0.07	-0.16	-0.06
p_3p_4	-0.01	-0.08	-0.18	-0.08	-0.08	-0.21
p_3p_5	-	-	-	-0.27	-0.26	-0.27
p_4p_5	-	-	-	-0.19	-0.24	-0.25

M4 and M5 - models using Legendre polynomials of orders four and five and including multiple (M) regressions of fixed age-season of calving lactation curves.

The trajectories of the additive genetic, permanent and temporary environment variance estimates for test-day milk yields showed a similar pattern for all lactations, Legendre polynomials of orders 4 and 5 and also for the alternatives of modeling the regression of fixed lactation curves (Figures 1 to 3).

Additive genetic variances estimates did not differ between single or multiple regression of fixed lactation curves models. The higher estimates were observed between 210 and 270 days of lactation, decreasing thereafter to the end of lactation (Figure 1).

The decrease in genetic variance in the extreme periods of lactation of the Holstein breed has been commonly observed in several studies (Pool et al., 2000, López-Romero & Carabaño, 2003, Costa et al., 2008). These authors have pointed that these trends may be associated to a lower number of production records and to the inadequacy of the mathematical functions to describe the environmental and genetic effects in these periods of lactation.

The permanent (Figure 2) and temporary environment (Figure 3) variance estimates obtained by single or multiple regressions of fixed lactation curves models were practically identical, differing only in their magnitudes across lactations.

Permanent environmental variances showed higher variability at the beginning and at the end of lactation (Figure 2). These results are similar to those observed by

Pool et al. (2000), Lopez-Romero & Carabaño (2003), Cobuci et al. (2005), Fujji & Suzuki (2006) and Araújo et al. (2006) for first lactation test-day milk yield of Holstein cows.

Figure 2 - Permanent environmental variance in the first (La1), the second (La2) and the third (La3) lactations obtained by random regression models with Legendre polynomial of order four (A) or five (B).
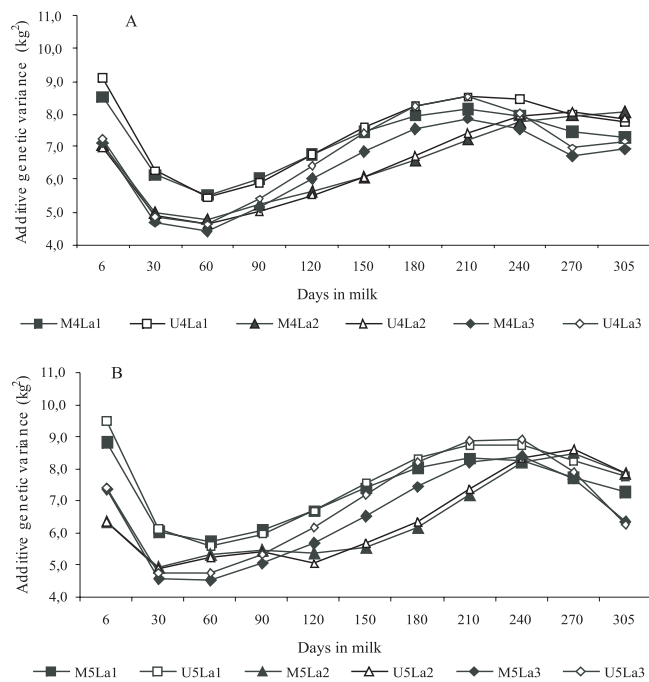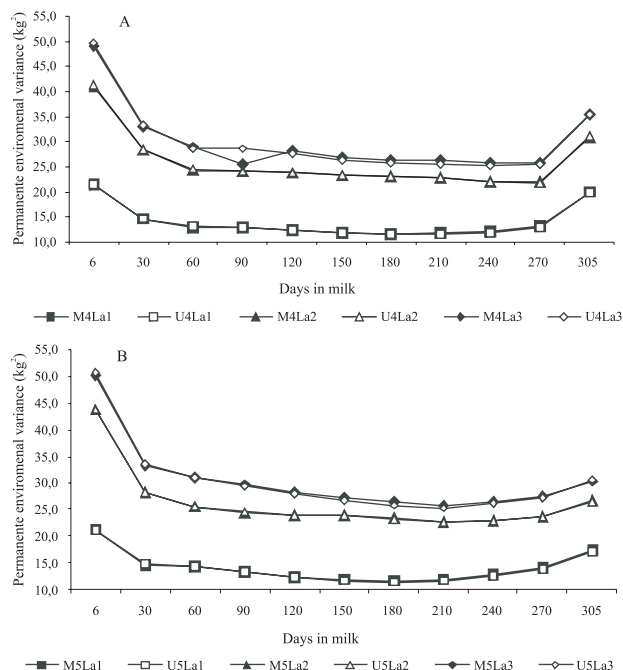

Figure 1 - Additive genetic variance in the first (La1), the second (La2) and the third (La3) lactations obtained by random regression models with Legendre polynomial of order four (A) or five (B).
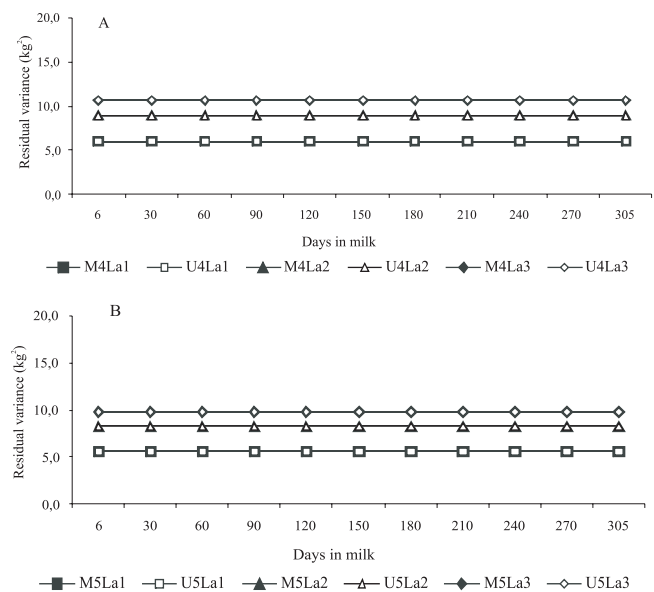

Figure 3 - Residual variance estimates in the first (La1), second (La2) and third (La3) lactations obtained by random regression models with Legendre polynomial of order four (A) or five (B).

Genetic parameters for milk production by using random regression models with ...

563

The temporary environment (residual) variance estimates increased from the first to third the lactation and showed similar magnitudes for single (U) or multiple (M) regression of fixed lactation curves models (Figure 3). Residual variances estimates decreased by approximately 7.5% when the order of the Legendre polynomial increased from four to five.

The estimates of residual variance for the first, the second and the third lactation were, respectively, 6.08, 8.89 and 10.59 kg^2, for models M4 and U4, and 5.61, 8.21 and 9.79 kg^2 for models M5 and U5. These values are very close to the values 7.16, 6.23 and 4.79 kg^2 obtained in Holstein cattle by Cobuci et al. (2005), Araujo et al. (2006) and Costa et al. (2008), respectively.

Costa et al. (2008) concluded that although the residual variance for test-day milk yield was heterogeneous across the lactation period, assuming that homogeneous residual variance would be a parsimonious option when making decisions about fitting test-day milk yields of Holstein cows in Brazil. However, El Faro & Albuquerque (2003) indicated that heterogeneous residual variances should be considered in modeling Caracu test-day milk yield by random regression. Fujii & Suzuki (2006) reported heterogeneity of residual variances over the years but concluded that by ignoring it did not affect the ranking of breeding values of the best bulls and cows of the Holstein breed in Japan.
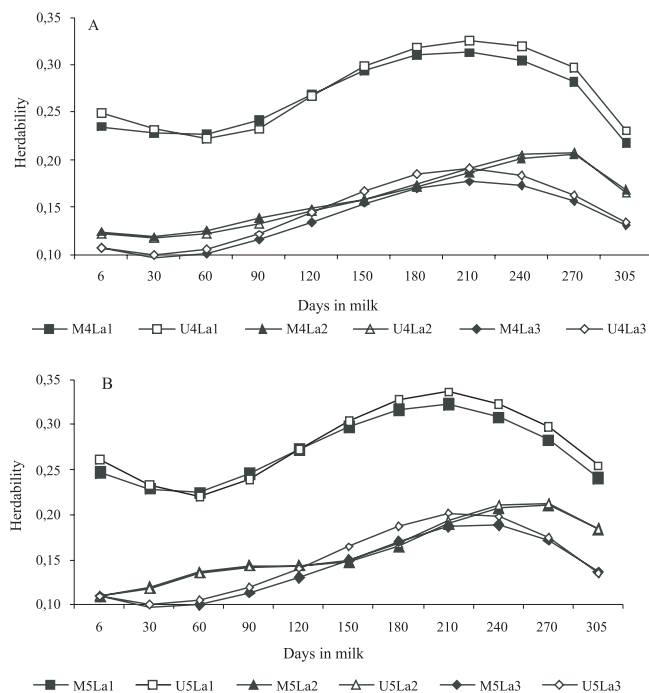

Figure 4 - Heritability estimates in the first (La1), second (La2) and third (La3) lactations obtained by random regression models with Legendre polynomial of order four (A) or five (B).

Heritability estimates practically did not differ between single and multiple regression of fixed lactation curves models but were larger for the first than for the second and third lactations (Figure 4). The small differences in heritability estimates between models (U4, U5 or M4 and M5) do not indicate a preferred order of the Legendre polynomial. This result does not agree with the reports of several authors that have emphasized that the order of Legendre polynomial in random regression models is an important aspect, inasmuch as the genetic parameter estimates may differ according to the order of this polynomial (Cobuci et al., 2006; Costa et al., 2008).

Heritability estimates obtained by models that considered multiple regressions for fixed lactation curves (M4 and M5) ranged from 0.22 to 0.32, from 0.11 to 0.21, and from 0.10 to 0.19, respectively, for the first, the second and the third lactations and from 0.23 to 0.34, from 0.11 to 0.21 and from 0.10 to 0.20 for the model that considered a single regression of the fixed curve of lactation (U4 and U5). Higher heritability values for the first lactation than for the following two lactations were also observed by Liu et al. (2000).

Trends of the heritability estimates during the second and the third lactation were more similar than those between the first and the second or the third lactation. In general, values increased from the beginning until 210-240 days of lactation and decreased thereafter to the end of the lactation. These trends are similar to those observed by Liu et al. (2000), Jakobsen et al. (2002), Bormann et al. (2003) and De Roos et al. (2004).

Variations in heritability estimates across lactation were associated to different trends in genetic and permanent environment variances. During lactation, while permanent environmental variance remained relatively constant (Figure 2), the additive genetic variability gradually increased (Figure 1), resulting in higher estimates of heritability in mid lactation than in the extreme periods. During late lactation, estimates of heritability decreased due to simultaneous reduction in genetic variance and increase in permanent environmental variance.

Similarly to the estimates of heritability, genetic correlation between milk production in selected days of lactation did not differ among lactation orders, regardless of the regression of fixed curve of lactation model and the Legendre polynomial used (Figures 5 and 6). As usually observed in most of studies using random regression models, genetic correlation between adjacent test-day were high, but decreased as the length between them increased.

The values of genetic correlations ranged from 0.13 to 0.99 for the first, from 0.16 to 0.99 for the second, and from
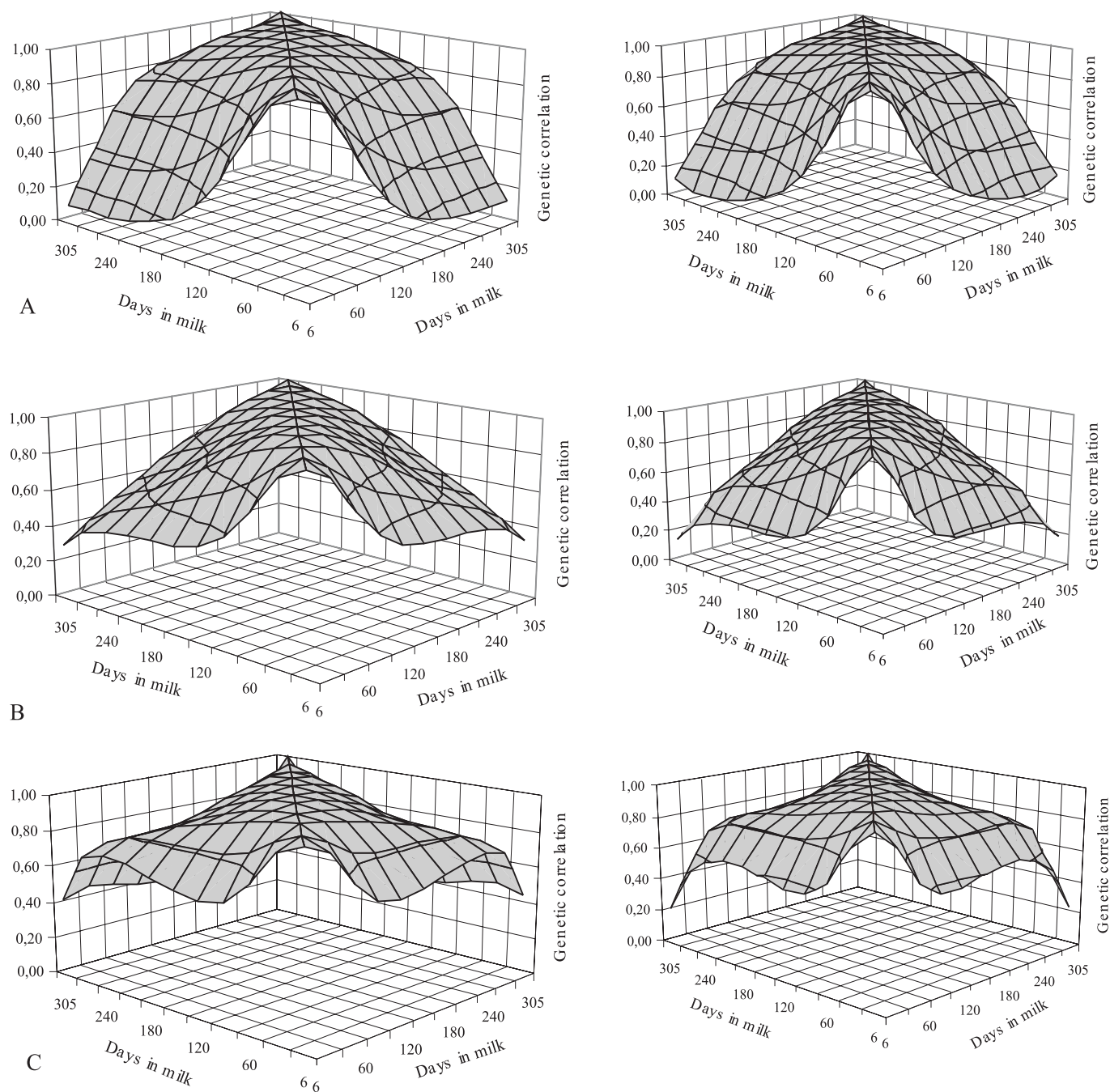
Figure 5 - Genetic correlations in the first (A), second (B) and third (C) lactations obtained by fitting single (left) or multiple (right) regressions models with Legendre polynomial of order four.

0.20 to 0.99 for the third lactation. The variation in genetic correlation estimates was larger in the first lactation than in the second and third lactations, although the trends within lactation were similar for all parities. These results are in agreement with previous studies which have reported the effect of parity on the estimation of genetic parameters in dairy cattle (Liu et al. 2000; Guo et al., 2002).

Jakobsen et al. (2002) and Araújo et al. (2006) reported genetic correlations estimates higher than 0.40 for first lactation test-day milk yield of Holstein cows, therefore much higher than some estimates observed in this study. However, lower estimates, even close to zero, were obtained for genetic correlations between test-day milk yields in extreme periods of lactation by Strabel & Misztal (1999) and Cobuci et al. (2005) when using different functions in random regression models.

The choice for the best random regression model has been commonly taken based on test results of Akaike

Genetic parameters for milk production by using random regression models with ...
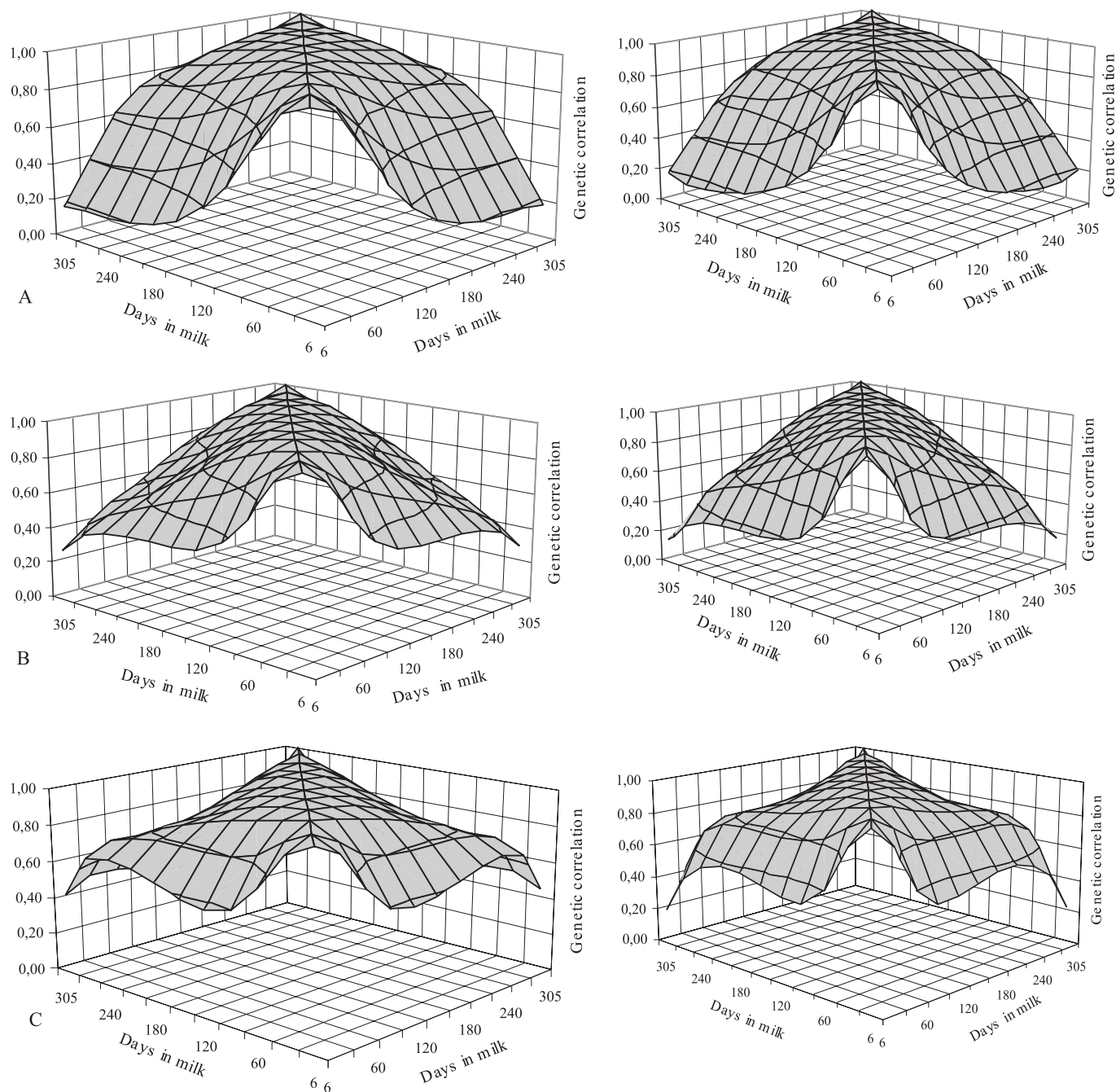
565

Figure 6 - Genetic correlations in the first (A), second (B) and third (C) lactations obtained by fitting single (left) or multiple (right) regressions models with Legendre polynomial of order five.

information criterion (AIC) or Bayesian information criterion (BIC). Both tests allow the comparison of non nested models and penalize the most parameterized models, the latter being more rigorous, with a tendency to penalize more, the most parsimonious models.

The values of -2 logL increased and AIC and BIC decreased as the order of Legendre polynomial increased in each lactation. Regardless of the order of the polynomial used, AIC and BIC values indicated that the models including

multiple regressions of fixed lactation curves better fitted test-day milk yield in the first lactation, but the single regression fixed lactation curve model as the best for fitting test-day milk yields in the second and third lactations (Table 6).

Misztal et al. (2000) observed that the choice of the mathematical function to describe fixed and random effects is the key element in fitting random regression models. After evaluating the Wilmink and Ali & Schaeffer parametric

Table 6 - Analysis of goodness of fit for different random regression models

Lactation	Model[1]	AIC[2]	BIC[3]	-2log(L)[4]	CM[5]	CL[6]
1	M4	1161274,70737	1161491,54145	1161232,70737	1	3
1	U4	1161569,86251	1161786,69658	1161527,86251	2	4
1	M5	1156629,94216	1156950,03056	1156567,94216	1	1
1	U5	1156850,25320	1157170,34160	1156788,25320	2	2
2	M4	559087,62203	559287,79099	559045,622035	2	4
2	U4	559037,52129	559237,69025	558995,521298	1	3
2	M5	557325,02251	557620,51002	557263,022517	2	2
2	U5	557202,47309	557497,96060	557140,473092	1	1
3	M4	203977,17673	204155,85322	203935,176732	2	4
3	U4	203972,92881	204151,60530	203930,928813	1	3
3	M5	203364,49922	203628,25975	203302,499220	2	2
3	U5	203332,53186	203596,29240	203270,531863	1	1

[1] M4, U4, U5 and M5 - models using Legendre polynomials of orders four and five and including single (U) or multiple (M) regression of fixed lactation curves; [2]AIC - Akaike information criterion; [3]BIC - Bayesian information criterion; [4]-2log (L) – value -2log of likelihood function (L); [5]CM – ranking order for the U and M models within the same Legendre polynomial order according to AIC and BIC; [6]CL – ranking order of U and M models, using Legendre polynomials of orders 4 and 5, within lactation, according to AIC and BIC.

functions and Legendre polynomials, Cobuci et al. (2006) and Costa et al. (2008) reported that the Legendre polynomials of orders 4 and 5 were the most appropriated for fitting test-day and persistency of milk yield, by random regression, in Holstein cows in Brazil.

Most of the studies do not indicate a consensus about the best criterion for choosing among different models and functions used to fit test-day records by random regression. According to Liu et al. (2006), as far as genetic evaluation essentially involves the use of computers, the final decision about a model may also be based on computational requirements. Besides this practical aspect, flexibility and robustness should also be considered when making decisions on the best random regression model to be used (Druet et al., 2003 and Cobuci et al., 2006).

The indication of a specific model to fit test-day milk yield has been difficult to be achieved. According to Kamidi (2005), this empirical state is likely to remain as long as modelled production patterns traverse combination of climate, management, health, age and parity order effects.

Conclusions

The use of multiple regressions for fixed lactation curves improves the goodness of fit of test day milk yield by random regression models without changing the magnitude of genetic parameters for milk yield along lactations of Holstein cows. The single regression should not be used to replace the multiple regression of fixed lactation curves model.

References

AKAIKE, H. **Information theory and an extension of the maximum likelihood principle**. In: INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY, 2., 1973, Busapeste. **Proceedings...** Budapeste, 1973. p.267-281.

ARAÚJO, C.V.; TORRES, R.A.; COSTA, C.N. et al. Uso de funções ortogonais para descrever a produção de leite no dia de controle por meio de modelos de regressão aleatória. **Revista Brasileira de Zootecnia**, v.35, p.967-974, 2006.

BORMANN, J.; WIGGANS, G.R.; DRUET, T. et al. Within-herd effects of age at test day and lactation stage on test-day yields. **Journal of Dairy Science**, v.86, p.3765-3774, 2003.

COBUCI, J.A.; COSTA, C.N.; TEIXEIRA, A.F. et al. Utilização dos polinômios de Legendre e da função de Wilmink em avaliações genéticas para persistência na lactação de animais da raça Holandesa. **Arquivo Brasileiro de Medicina Veterinária e Zootecnia**, v.58, n.4, p.614-623, 2006.

COBUCI, J.A; EUCLYDES, R.F; LOPES, P.S. et al. Genetic parameters for test-day milk yield in Holstein cows, estimated by random regression models. **Genetic and Molecular Biology**, v.28, p.75-83, 2005.

COSTA, C.N.; MELO, C.M.R.; PACKER, I.U. et al. Genetic parameters for test day milk yield of first lactation Holstein cows estimated by random regression using Legendre polynomials. **Revista Brasileira de Zootecnia**, v.37, n.4, p.602-608, 2008.

DE ROOS, A.P.W.; HARBERS, A.G.F.; DE JONG, G. Random herd curves in a test-day model for milk, fat, and protein production of dairy cattle in the Netherlands. **Journal of Dairy Science**, v.87, p.2693-2701, 2004.

DRUET, T.; JAFFRÉZIC, F.; BOICHARD, D. et al. Modeling lactation curves and estimation of genetic parameters for first lactation test-day records or french holstein cows. **Journal of Dairy Science**, v.86, p.2480-2490, 2003.

EL FARO, L.; ALBUQUERQUE, L.G. Utilização de modelos de regressão aleatória para produção de leite no dia do controle, com diferentes estruturas de variâncias residuais. **Revista Brasileira de Zootecnia**, v.32, n.5, p.1104-1113, 2003.

FUJII, C.; SUZUKI, M. Comparison of homogeneity and heterogeneity of residual variance using random regression test-day models for first lactation japanese Holstein cows. **Animal Science Journal**, v.77, p.28-32, 2006.

GUO, Z.; LUND, M.S.; MADSEN, P. Genetic parameter estimation for milk yield over multiple parities and various lengths of lactation in Danish Jerseys by random regressions models. **Journal of Dairy Science**, v.85, n.6, p.1596-1606, 2002.

JAKOBSEN, J.H.; MADSEN, P.; JENSEN, J. et al. Genetic parameters for milk production and persistency for Danish

Genetic parameters for milk production by using random regression models with ...

567

Holstein estimated in random regression models using REML. **Journal of Dairy Science**, v.85, n.6, p.1607-1616, 2002.

KAMIDI, R.E. A parametric measure of lactation persistency in dairy cattle. **Livestock Production Science**, v.96, p.141-148, 2005.

LIU, Y.X.; ZHANG, J.; SCHAEFFER L.R. et al. Short communication: Optimal random regression models for milk production in dairy cattle. **Journal of Dairy Science**, v.89, p.2233-2235, 2006.

LIU, Z.; REINHARDT, F.; REENTS, R. Estimating parameters of a random regression test day model for fist three lactation milk production traits using the covariance function approach. **Interbull Bulletin**, v.25, p.74-80, 2000.

LÓPEZ-ROMERO, P.; CARABAÑO, M.J. Comparing alternative random regression models to analyze first lactation daily milk yield data in Holstein-Friesian cattle. **Livestock Production Science**, v.82, p.81-96, 2003.

MISZTAL, I. [2005]. REMLF90 manual. Disponível em: <ftp://nce.ads.uga.edu/pub/ignacy/blupf90/>. Acesso em: 11/10/2005.

MISZTAL, I.; STRABEL, T.; JAMROZIK, J. et al. Strategies for estimating the parameters needed for different test-day models. **Journal of Dairy Science**, v.83, p.1125-1134, 2000.

MUIR, B.L.; KISTEMAKER, G.; JAMROZIK, J. et al. Genetic parameters for a multiple-trait multiple-lactação random regression test-day model in Italian holsteins. **American Dairy Science Association**, v.90, p.1564-1574, 2007.

POOL, M.H.; JANSS, L.L.G.; MEUWISSEN, T.H.E. Genetic parameters of Legendre polynomials for first parity lactation curves. **Journal of Dairy Science**, v.83, p.2640-2649, 2000.

SCHAEFFER, L.R. Application of random regression models in animal breeding. **Livestock Production Science**, v.86, p.35-45, 2004.

STRABEL, T.; JAMROZIK, J. Genetic analysis of milk production traits of Polish black and white cattle using large-scale random regression test-day models. **American Dairy Science Association**, v.89, p.3152-3163, 2006.

SCHWARZ, G. Estimating the dimension of a model. **The Annals of Statistics**, v.6, n.2, p.461-464, 1978.

STRABEL, T.; MISZTAL, I. Genetic parameters for first and second lactation milk yields of Polish Black and White cattle with random regression test-day models. **Journal of Dairy Science**, v.83, p.1125-1134, 2000.

STRABEL, T.; PTAK, E.; SZYDA, J. et al. Multiple-lactation random regression test-day model for Polish Black and White cattle. **Interbull Bulletin**, v.32, p.133-136, 2004.