Animal (2012), **6:2**, pp 212–220 © The Animal Consortium 2011
doi:10.1017/S1751731111001534

Random regression analyses using B-spline functions to model growth of Nellore cattle

A. A. Boligon[1†], M. E. Z. Mercadante[2], R. B. Lôbo[3], F. Baldi[1] and L. G. Albuquerque[1]

[1]Departamento de Zootecnia, Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal, CEP 14884-900, São Paulo, Brazil; [2]Instituto de Zootecnia, Estação Experimental de Zootecnia de Sertãozinho, 14160-000 Sertãozinho, São Paulo, Brazil; [3]Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, 14049-900 Ribeirão Preto, São Paulo, Brazil

(Received 5 April 2011; Accepted 24 June 2011; First published online 12 September 2011)

The objective of this study was to estimate (co)variance components using random regression on B-spline functions to weight records obtained from birth to adulthood. A total of 82 064 weight records of 8145 females obtained from the data bank of the Nellore Breeding Program (PMGRN/Nellore Brazil) which started in 1987, were used. The models included direct additive and maternal genetic effects and animal and maternal permanent environmental effects as random. Contemporary group and dam age at calving (linear and quadratic effect) were included as fixed effects, and orthogonal Legendre polynomials of age (cubic regression) were considered as random covariate. The random effects were modeled using B-spline functions considering linear, quadratic and cubic polynomials for each individual segment. Residual variances were grouped in five age classes. Direct additive genetic and animal permanent environmental effects were modeled using up to seven knots (six segments). A single segment with two knots at the end points of the curve was used for the estimation of maternal genetic and maternal permanent environmental effects. A total of 15 models were studied, with the number of parameters ranging from 17 to 81. The models that used B-splines were compared with multi-trait analyses with nine weight traits and to a random regression model that used orthogonal Legendre polynomials. A model fitting quadratic B-splines, with four knots or three segments for direct additive genetic effect and animal permanent environmental effect and two knots for maternal additive genetic effect and maternal permanent environmental effect, was the most appropriate and parsimonious model to describe the covariance structure of the data. Selection for higher weight, such as at young ages, should be performed taking into account an increase in mature cow weight. Particularly, this is important in most of Nellore beef cattle production systems, where the cow herd is maintained on range conditions. There is limited modification of the growth curve of Nellore cattle with respect to the aim of selecting them for rapid growth at young ages while maintaining constant adult weight.

Keywords: growth curve, covariance functions, segmented polynomials, Nellore cattle

Implications

Random regression models (RRMs) applying spline functions could be an option to model weights until mature age, as these functions are less susceptible to the problems frequently observed with orthogonal Legendre polynomials. However, to date there are no studies that have applied RRM on B-spline functions to estimate genetic parameters for weights in Nellore cattle. The results of this research showed that RRMs using B-spline functions revealed similar description of the trajectory of the variances and heritability estimates. Moreover, random regression on B-spline functions required lower computational requirement and faster convergence in relation to Legendre polynomials or multi-trait analysis. In genetic evaluation programs, with large data set, the application of RRMs using B-spline functions could still be advantageous compared with Legendre polynomials.

Introduction

Random regression models (RRMs) based on Legendre polynomials have been used to study genetic parameters for growth traits in beef cattle (Albuquerque and Meyer, 2001; Nobre *et al.*, 2003; Arango *et al.*, 2004; Meyer, 2004, Nephawe, 2004; Boligon *et al.*, 2010). These models present the advantage of using all records of the animal and of eliminating the need for the correction of weights for standard ages. In addition, it provides more accurate estimates of breeding values for any age within the interval of the records considered when compared with conventional multi-trait models (Tier and Meyer, 2004).

[†] E-mail: arioneboligon@yahoo.com.br

Over the past few years, various researchers have pointed out that the use of random regression analysis based on Legendre polynomials to model weights in beef cattle may present some undesired properties, such as the need for employing high-degree polynomials and large increases of (co)variances along the curve (Meyer, 1999; Nobre *et al.*, 2003; Sakaguti *et al.*, 2003), especially at ages with a small number of weight records.

According to Meyer (2005a), analysis of weight records in beef cattle using RRMs with cubic, quadratic or even higher order polynomials has resulted in erratic and implausible estimates of variance components and genetic parameters. These problems are mainly observed in the case of data sets that contain a small number of records for the last ages, as adult weights, associated with a small number of records per animal, and in the case of analyses in which the number of records available for the animals is smaller than the order of the fitted polynomials. In this sense, Boligon *et al.* (2010), using Legendre polynomials to model changes in variances with age for weights from birth to adulthood, reported convergence problems with the increase in polynomial order.

In this respect, RRMs applying spline function offer a viable method to model weight records until mature age, as these functions are less susceptible to the problems frequently observed with orthogonal Legendre polynomials. According to Meyer (2000), spline function permitted the fitting of lower-degree polynomials in short segments of the growth trajectory, thus resulting in greater flexibility when fitting seasonal fluctuations in cattle growth.

One particular type of spline function is called B-splines, which are preferred because of their excellent numerical properties (Ruppert *et al.*, 2003). According to Rice and Wu (2001), B-splines are adequate to model random effects in mixed model analyses and are also efficient in estimating covariance functions.

In beef cattle, spline functions have been used to describe genetic variations in the growth of Angus cattle using weights from birth to 820 days of age (Meyer, 2005b), in Gelbvieh from birth to 1 year of age (Iwaisaki *et al.*, 2005) and in Polled Hereford dams from 1.5 to 7 years of age (Meyer, 2005c). Similarly, Sánchez *et al.* (2008) applied linear spline RRMs to analyze weights of crossbred animals from birth to 1 year of age and showed that these models can also be used for genetic evaluation of multi-breed populations.

In Brazil, Baldi *et al.* (2010) compared estimates of (co)variance functions for weights of Canchim females from birth to adulthood obtained by random regression using Legendre polynomials or B-splines. The results showed that B-spline functions provided a better description of the trajectory of variance estimates than Legendre polynomials, especially at older ages. However, there are no reported studies in the literature applying RRMs on B-spline functions to estimate genetic parameters for weight records in Nellore cattle.

Most Brazilian beef cattle are raised on range conditions, characterized by high temperatures and low-quality pastures. Of late, there have been concerns about consequences of selection to increase weights at young ages on cow size, which could result in higher costs to maintain the cow herd. In addition, for low-input environment, medium-sized cows seem to be advocated (Jenkins and Ferrell, 1994). In this scenario, to maintain desirable adult size it is necessary to consider adult weight in beef cattle selection indices. However, there are difficulties to include adult weights in genetic evaluations because of the scarcity of weight records at later ages and the lack of adequate models to estimate (co)variance components at mature ages. Apparently, RRMs on B-spline functions offer advantages over the Legendre polynomials and the traditional finite dimensional models (Tier and Meyer, 2004) and could be used for genetic evaluation of growth traits.

The objective of this study was to estimate covariance functions for direct and maternal additive genetic effects, animal and maternal permanent environmental effects, and subsequently, to obtain genetic parameters for growth traits in Nellore cattle, from birth to adulthood, using RRMs on B-splines of animal age.

Material and methods

Data

The data used in this study were obtained from eight herds located in the State of São Paulo that are part of the Nellore Breeding Program (PMGRN/Nellore Brasil), where a genetic evaluation program started in 1987. The animals were weighed at intervals of 45 days from birth to yearling, and those remaining in the herds as seed stock continue to be weighed routinely at this interval. In the case of the data set used here, births occurred throughout the year, with a larger number of births during Spring (September to November) and Summer (December to February). The animals were weaned on average at 240 days of age. Only weight records of animals reared on pasture without supplementation and nursed by their biological mothers were included in the analysis. Age of dam at calving ranged from 2 to 24 years and 13 age classes were considered for analysis, because the last class included dams aged 15 years or older at calving.

The contemporary group for records from birth to 3 years of age was defined including animals born on the same farm, year and season of birth, and belonging to the same group for age at recording (birth and every 45 days of age). Consequently, the variation in age within the contemporary group was less than 45 days. After 3 years of age, the contemporary group consisted of farm, year and season of birth and year and season of recording. Weight records outside the intervals given by the mean of the contemporary group ± 3 s.d. were excluded. Contemporary groups containing fewer than eight animals were excluded, for a total of 2693 contemporary groups.

Animals with fewer than five weight records were excluded. The mean number of records per animal was 10.07, with a maximum of 43. There were 813, 2133, 1243, 1138, 528, 282 and 2008 animals with 5, 6, 7, 8, 9, 10 and 11 to 43 weight records, respectively, corresponding to 63.83%, 48.57% and 34.60% of animals with at least 7, 8 and 9 records, respectively.

The data set analyzed consisted of 81 712 weight records from birth to 8 years of age, comprising 8144 females, daughters to 470 sires and 5185 cows. The relationship matrix was constructed using information from all generations available and contained 12 596 animals.

Analysis

All models included direct and maternal genetic and animal and maternal permanent environmental effects as random effects, in addition to residual effects. The random effects were modeled using B-spline functions (Marx and Eiler, 1998). Contemporary group and linear and quadratic effects of dam age at calving within the age class of the animal (five classes: 2 to 3, 4 to 5, 6 to 8, 9 to 11, 12 to 24 years of age) were included as fixed effects.

A fourth-order orthogonal Legendre polynomial was used to model the population's mean growth trajectory. Residual variances were modeled considering five age classes (1, 2 to 240, 241 to 550, 551 to 730 and 731 to 2889 days). The matrix representation of the models is

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_1 a} + \mathbf{Z_2 m} + \mathbf{W_1\, pe} + \mathbf{W_2\, mpe} + \mathbf{e},$$

where **y** is the vector of observations, **b** is the vector of fixed effects, **a** is the vector of random coefficients for additive direct effects, **m** is the vector of random coefficients for additive maternal effects, **pe** is the vector of random coefficients for animal permanent environmental effects, **mpe** is the vector of random coefficients for maternal permanent environmental effects, **e** is the vector of residual effects and $\mathbf{X}$, $\mathbf{Z_1}$, $\mathbf{Z_2}$, $\mathbf{W_1}$ and $\mathbf{W_2}$ are the corresponding incidence matrices. The model is based on the following assumptions:

$$E\begin{bmatrix}\mathbf{y}\\ \mathbf{a}\\ \mathbf{m}\\ \mathbf{pe}\\ \mathbf{mpe}\end{bmatrix} = \begin{bmatrix}\mathbf{Xb}\\ 0\\ 0\\ 0\\ 0\end{bmatrix};\quad V\begin{bmatrix}\mathbf{a}\\ \mathbf{m}\\ \mathbf{pe}\\ \mathbf{mpe}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{K_a}\otimes\mathbf{A} & 0 & 0 & 0 & 0\\ 0 & \mathbf{K_m}\otimes\mathbf{A} & 0 & 0 & 0\\ 0 & 0 & \mathbf{K_{pe}}\otimes\mathbf{I}_{N_A} & 0 & 0\\ 0 & 0 & 0 & \mathbf{K_{mpe}}\otimes\mathbf{I}_{N_M} & 0\\ 0 & 0 & 0 & 0 & \mathbf{R}\end{bmatrix},$$

where $\mathbf{K}_a$, $\mathbf{K}_m$, $\mathbf{K}_{pe}$ and $\mathbf{K}_{mpe}$ are (co)variance matrices between random regression coefficients for additive direct, additive maternal and animal and maternal permanent environmental effects, respectively, **A** is the relationship matrix, **I** is an identity matrix, N_A is the number of animals for which records are available, N_M is the number of dams with progeny records, ⊗ is the Kronecker product between matrices and **R** is a block diagonal matrix containing residual variances. Correlations between random regression coefficients for different effects were set to zero.

The B-spline function can be defined recursively (Meyer, 2005b), with basis functions of degree $p = 0$ assuming values of unity for all points in a given interval; otherwise, the value is zero. For the kth interval given by knots T_k and $T_k < T_k + 1$,

$$B_{k,0}(t) = \begin{cases} 1 & \text{if } T_k \le t < T_k + 1,\\ \text{otherwise } 0. \end{cases}$$

High-degree basis functions, $B_{k,p}$ for $p > 0$, are then determined from the values of the lower degree basis functions, already evaluated, and the width of the adjoining intervals between knots (Meyer, 2005b). According to De Boor (2001), the function can be described as

$$B_{k,p}(t) = \frac{t - T_k}{T_{k+p} - T_k} B_{k,p} 1^{(t)} + \frac{T_{k+p+1} - 1}{T_{k+p+1} - T_{k+1}} B_{k+1,p} 1^{(t)}.$$

Linear (L), quadratic (Q) and cubic (C) polynomials were considered for each individual segment, with basis functions of degree $p = 1$, 2 and 3, respectively. Polynomials of the same degree were included in the model for all random effects. The m knots were chosen to divide the age at recording into $m-1$ equally spaced intervals and the external knots are located in classes 1 and 2889 for all models. Up to seven knots (six segments) were used for direct additive genetic and permanent environmental effects of the animal, with the same number of knots fitted for both effects. A single segment with two knots at the end points of the curve was used for maternal genetic and permanent environmental effects.

The number of random regression coefficients to model the trajectory of the linear, quadratic and cubic basis functions is given by m, $m + 1$ and $m + 2$, respectively. The nomenclature of the RRMs follows the pattern 'BSX k_a, k_m, k_{pe}, k_{mpe}', where $X =$ L, Q or C, that is, the degree of the polynomial segment, and k_a, k_m, k_{pe} and k_{mpe} are the specific numbers of random regression coefficients for direct additive genetic, maternal genetic and animal and maternal permanent environmental effects, respectively. Thus, model BSL3232 consists of a linear B-spline function with three random regression coefficients for direct additive genetic and permanent environmental effects of the animal, and two random regression coefficients for maternal genetic and permanent environmental effects.

The covariance components and genetic parameters of the models were estimated by the restricted maximum likelihood method using the Wombat statistical program (Meyer, 2006). Direct additive genetic variances for weights at different ages ($\hat{\sigma}^2_{a_i}$) were obtained using

$$\hat{\sigma}^2_{a_i} = f_i' \Lambda_A f_i,$$

where $\mathbf{f}_i'$ is the vector of covariates corresponding to the B-spline function of age at recording in days, and Λ_A is the matrix of direct additive genetic covariances between random regression coefficients. Maternal genetic, animal and maternal permanent environmental variances were obtained in a similar manner replacing Λ_A by Λ_M, Λ_{PE} or Λ_{MPE}, respectively. The Λ_M, Λ_{PE} and Λ_{MPE} are the matrices of maternal genetic, animal and maternal permanent environmental covariances between random regression coefficients, respectively. The residual variances ($\hat{\sigma}^2_{e_i}$) were considered to change with age (described by five age classes). Thus, direct heritabilities of weights at different ages ($h^2_{a_i}$) were calculated using the equation:

$$h^2_{a_i} = \hat{\sigma}^2_{a_i} / (\hat{\sigma}^2_{a_i} + \hat{\sigma}^2_{m_i} + \hat{\sigma}^2_{pe_i} + \hat{\sigma}^2_{mpe_i} + \hat{\sigma}^2_{e_i}).$$

The same equation shown above was used to estimate maternal heritability ($h^2_{\mathrm{m}_i}$) and animal (pe^2_i) and maternal (mpe^2_i) permanent environmental variances as a proportion of total phenotypic variances, replacing the numerator of this equation by ($\hat{\sigma}^2_{\mathrm{m}_i}$), ($\hat{\sigma}^2_{\mathrm{pe}_i}$) or ($\hat{\sigma}^2_{\mathrm{mpe}_i}$), respectively (for details see Meyer, 2001; Mrode, 2005).

The models were compared using the Akaike information criterion (AIC) and the Schwarz–Bayesian information criterion (BIC; Schwarz, 1978), and by the analysis of variances and correlations estimated for weights at the different ages considered. The AIC and BIC considered the lack of fit and the number of parameters estimated in the model. The AIC can be used to compare models including the same fixed effects but different variance structures. The judgment is the same in the case of the BIC, but this criterion attributes a higher penalty to models containing a larger number of parameters than the AIC. The information criteria are calculated as follows:

$$\mathrm{AIC} = -2\log\mathrm{L} + 2p,$$
$$\mathrm{BIC} = -2\log\mathrm{L} + p\log(N - r),$$

where p is the number of parameters in the model, N is the total number of observations, r is the rank of the incidence matrix of fixed effects in the model and logL is the logarithm of the likelihood function. Lower values for AIC and BIC indicate a better fit.

The results obtained with the B-spline functions were compared with those obtained using random regression on fourth- and sixth-order orthogonal Legendre polynomials for direct additive genetic and permanent environmental effects of the animal, respectively, and on third-order polynomials for maternal genetic and permanent environmental effects (Leg4363) as proposed by Boligon *et al.* (2010). The same fixed effects were considered for all RRM.

A multi-trait analysis was also performed using a finite dimension model for weights at birth, weaning, yearling, 18 months of age and 2, 3, 4, 5 and 6 years of age. The model for all traits included direct additive and maternal genetic effects, maternal permanent environmental effects and residual effects as random effects. Contemporary group (farm, year and season of birth) and linear and quadratic effects of animal age at recording (except for weight at birth) and dam age at calving were included as fixed effects. The covariance between direct and maternal genetic effects was set to zero in line with the suggestion of Albuquerque and Meyer (2001) and Meyer (2005b). Moreover, in Brazil this covariance is assumed to be zero in most of beef cattle evaluation programs. For more details, the multi-trait analysis applied was reported by Boligon *et al.* (2009), using the same data set.

Results and discussion

Weights increased with age, with a decrease in the growth rate after weaning, and a practically linear growth rate after 800 days of age (Figure 1). A marked decline in the number of animals was observed after 600 days of age, a finding suggesting that many farms only weigh females until they enter the first breeding season.

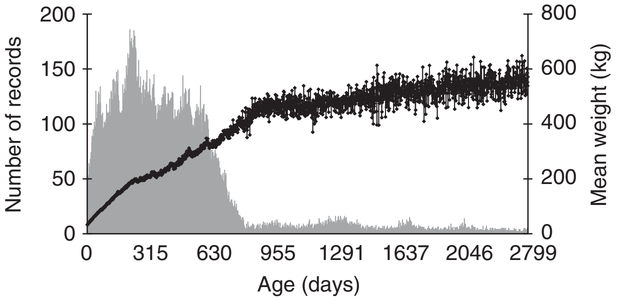

Figure 1 Number of records (bars) and mean weight (line) according to age.

The two different information criteria to compare the models (AIC and BIC) indicated the same models, and thus only the BIC results were shown. Table 1 summarizes the results of the model comparisons using the logarithm of the likelihood function (logL), and BIC obtained for the models used. The number of parameters ranged from 17 to 81 in the models studied. The logarithm of the likelihood function (logL) improved as the number of parameters in the models increased, irrespective of the degree of the function. Similar results were reported by Meyer (2005b) and Baldi *et al.* (2010) using B-splines (linear, quadratic and cubic) to model the growth trajectory of Angus cattle from birth to 821 days of age and with Canchim females from birth to 8 years of age, respectively.

None of the RRMs on B-splines as base functions with similar or lower number of parameters than the Leg4363 model had fitted better (Table 1). Therefore, applying RRMs on B-spline functions, a higher number of parameters were needed to reach a similar fit as Leg4363 model (Table 1). These results were in agreement with those reported by Meyer (2005a), that even with relatively few knots RRMs on B-spline functions tend to involve a larger number of coefficients than models fitting Legendre polynomials.

The logarithm of the likelihood function and the BIC results indicated that BSC7474 model, with 81 parameters, fitted the data best. However, this model presented implausible phenotypic variance estimates at extreme ages, compared with multi-trait and Leg4363 models (Figure 2). Alternatively, the model BSQ5353, with 47 parameters, showed variance component estimates similar to those from Leg4363 and multi-trait analysis. Increasing the number of knots or the polynomial order improved BIC; however, the variance estimates showed an unexpected behavior for later weights (not shown). Meyer (2005b) also reported high and implausible residual and phenotypic variances at more advanced ages when weights were fitted using a cubic B-spline function. It should be emphasized that in the case of B-spline functions a good fit of the data depends, in addition to the order of the individual segments, on the data structure, with problems mainly being observed when the data are distributed irregularly.

Table 1 *Models, number of parameters (n) in the model, position of knots, number of segments (s), likelihood function (log), and Schwarz–Bayesian information criterion (BIC) obtained for random regression models using B-splines and Legendre polynomials*

		Knots				
Model	*n*	Direct	Maternal	*s*	Log	BIC
BSL2222	17	1, 2889	1, 2889	1	−259 921	520 034
BSL3232	23	1, 1445, 2889	1, 2889	2	−254 460	509 179
BSL4242	31	1, 964, 1927, 2889	1, 2889	3	−245 791	491 932
BSL5252	41	1, 723, 1445, 2172, 2889	1, 2889	4	−231 487	463 437
BSL6262	53	1, 578, 1156, 1734, 2314, 2889	1, 2889	5	−224 011	448 620
BSL7272	67	1, 483, 964, 1445, 1927, 2409, 2889	1, 2889	6	−219 917	440 590
BSQ3333	29	1, 2889	1, 2889	1	−249 946	500 219
BSQ4343	37	1, 1444, 2889	1, 2889	2	−234 460	469 337
BSQ5353	47	1, 963, 1926, 2889	1, 2889	3	−228 741	458 012
BSQ6363	59	1, 723, 1445, 2172, 2889	1, 2889	4	−220 934	442 534
BSQ7373	73	1, 579, 1156, 1734, 2311, 2889	1, 2889	5	−218 714	438 252
BSC4444	45	1, 2889	1, 2889	1	−225 146	450 800
BSC5454	55	1, 1425, 2889	1, 2889	2	−222 349	445 319
BSC6464	67	1, 950, 1914, 2889	1, 2889	3	−219 045	438 846
BSC7474	81	1, 713, 1435, 2157, 2889	1, 2889	4	−215 107	431 128
		Legendre polynomials				
Leg4363[1]	48				−221 382	443 305
		Finite dimension model				
Multi-trait[2]	96				−215 096	431 247

[1] Legendre polynomials: 4, 3, 6 and 3 correspond to the order of the regression coefficients for direct additive and maternal genetic effects and animal and maternal permanent environmental effects, respectively.

[2] Weights at birth, weaning, yearling, 18 months of age and 2, 3, 4, 5 and 6 years of age.

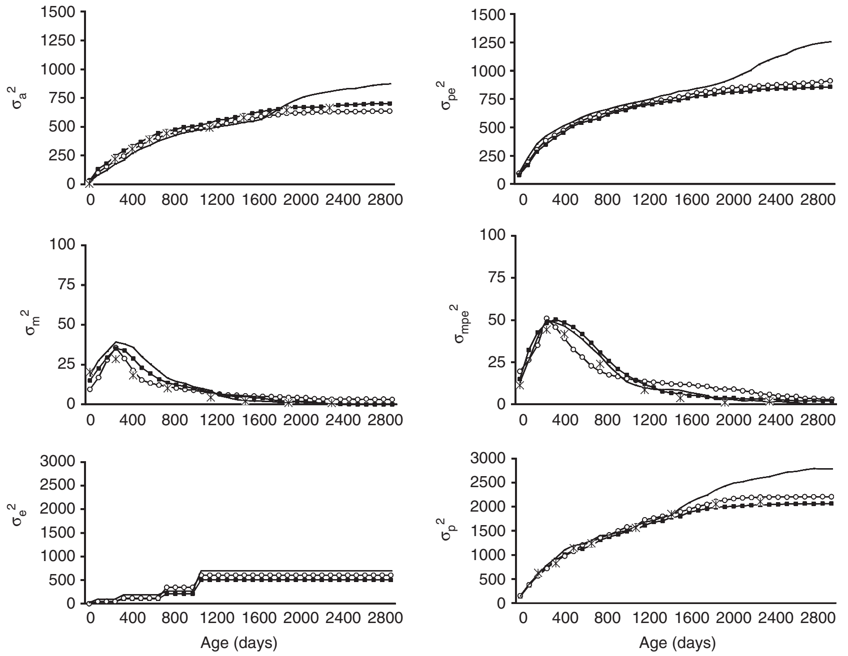

Figure 2 Estimates of direct additive genetic (σ_a^2), maternal genetic (σ_m^2), animal permanent environment (σ_{pe}^2), maternal permanent environment (σ_{mpe}^2), residual (σ_e^2) and phenotypic (σ_p^2) variance components obtained with models BSC7474 (—), BSQ5353 (■), Leg4363 (○) and by multi-trait analysis (✻).

The direct additive genetic variances obtained by BSC7474, BSQ5353, Leg4363 and multi-trait models were similar until almost 5 years of age (Figure 2). However, after this point, the additive direct variance estimates obtained with the BSC7474 model were higher than those estimated with BSQ5353 with Leg4363 models. Similar results were reported by

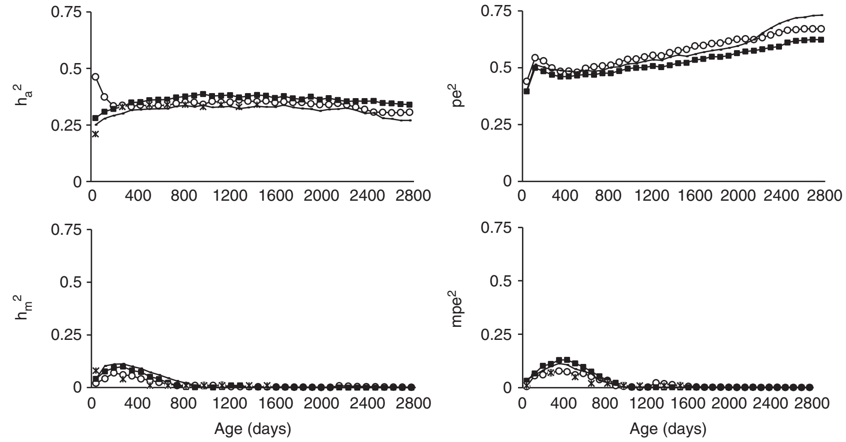

Figure 3 Estimates of direct (h_a^2) and maternal heritability (h_m^2) and estimates of animal (pe^2) and maternal (mpe^2) permanent environmental variance as a proportion of total phenotypic variance obtained with models BSC7474 (—), BSQ5353 (■) and Leg4363 (○) and by multi-trait analysis (⁎).

Baldi *et al.* (2010) with weight records from birth to mature age in cows of Canchim breed. These authors reported that for weights measured after 6 years of age, RRMs on B-spline functions fitting cubic regression showed lower additive direct variance estimates than those obtained with B-spline functions fitting quadratic regression or Legendre polynomials.

Using weight records of Gelbvieh cattle from birth to 365 days of age, Iwaisaki *et al.* (2005) reported similar direct additive genetic variance estimates obtained by random regression on linear spline functions or by multi-trait analysis (weights at birth, weaning and 1 year of age). For ages with enough number of weight records, no important differences in variance component estimates are expected when random regression on B-spline or Legendre polynomials as base functions were applied (Meyer, 2005b).

Similar animal permanent environmental variance estimates were obtained for all RRMs until almost 5 years of age. After this age, the BSQ5353 showed lower animal permanent environmental variance estimates than those obtained with BSC7474 (Figure 2). Similar to the studies of Meyer (2005b) and Baldi *et al.* (2010), the use of quadratic B-spline functions permitted a better control of the variances of animal permanent environmental effects compared with fits obtained with cubic B-spline functions and Legendre polynomials.

The maternal genetic and permanent environmental variance estimates with BSC7474, BSQ5353 and using Legendre polynomials (Leg4363) or multi-trait analysis showed similar trend (Figure 2). For these models, higher variances were estimated for weights close to 240 days of age, followed by a decline after this age and variances of zero at older ages. Similar to this study, Albuquerque and Meyer (2001) observed that maternal genetic variance estimates increased until 150 days of age in Nellore cattle and remained practically constant until 240 days of age, followed by a decrease thereafter.

Except for the beginning of the study period (first 60 days), the heritability estimates for direct additive genetic effects obtained using the models BSC7474, BSQ5353 and Leg4363 and by multi-trait analysis showed the same trend (Figure 3). For all ages, direct heritability estimates obtained by BSC7474 model were slightly lower than those obtained with other models. Direct heritability estimates obtained by multi-trait analysis ranged from 0.23 (for weight at birth) to 0.35 (for weight at 5 and 6 years of age). In contrast, heritabilities obtained with model BSQ5353 presented an increase from birth (0.28) to yearling (0.38) and remained practically constant until 8 years of age (0.35). Using random regression on B-spline functions, Meyer (2005b) reported higher direct heritabilities when compared with those obtained by single-trait analysis, especially for weights from weaning to 365 days of age. On the other hand, Iwaisaki *et al.* (2005) found higher direct heritability estimates by multi-trait analysis for weights at birth, weaning and 1 year of age (0.52, 0.36 and 0.59, respectively) compared with those obtained using linear spline functions (0.51, 0.28 and 0.48, respectively).

In Brazil, using random regression on B-spline functions for weights of Canchim females from birth to 8 years of age, Baldi *et al.* (2010) observed an increase in the direct heritability estimates from birth to 450 days of age, followed by a small decline, remaining practically constant after 3 years of age.

Except at the end of the period studied, the estimates of animal permanent environmental variances as a proportion of total phenotypic variance were similar applying random regression on B-splines (BSC7474 and BSQ5353) or models fitting Legendre polynomials (Figure 3). For weights obtained after 5 years of age, the BSC7474 model showed a slight increase in animal permanent environmental variance as a proportion of phenotypic variance. Using the BSQ5353 model, the animal permanent environmental variances as a proportion of phenotypic variance increased from birth (0.39) to 120 days of age (0.50), presented a small decrease at about 300 days of age and then tended to increase slightly until the end of the study period. Similarly, Baldi *et al.* (2010) reported an increase in animal permanent environmental variance as a proportion of phenotypic variance at the end of the study period. However, in contrast to this study, wider oscillations at the beginning of the study period were reported by these authors. According to Meyer (2005b), B-spline functions are less susceptible to end-of-range problems, especially for animal permanent environmental effects, than orthogonal polynomials when modeling the growth curve of beef cattle.

For weight at birth, the maternal genetic heritabilities estimated by RRMs using B-spline functions or Legendre polynomials were lower than those obtained by multi-trait analysis (Figure 3). After birth, the maternal heritability estimates obtained by random regression on B-splines (BSC7474 and BSQ5353) increased until close to weaning (240 days of age) and decreased thereafter. Values of practically zero were observed after 365 days of age. In contrast to the present findings, Meyer (2005b) and Iwaisaki *et al.* (2005) observed lower maternal genetic heritabilities when using spline functions compared with those obtained using a finite dimension model (single- or multi-trait analysis).

Maternal permanent environmental variances as a proportion of phenotypic variance estimated with the RRMs using B-spline functions followed the same trend as those obtained by multi-trait analysis (Figure 3). In contrast, the estimates obtained by random regression on Legendre polynomials were generally lower and close to zero. Meyer (2005b) using RRMs on B-splines reported lower maternal permanent environmental variance estimates as proportions of phenotypic variances than those obtained with finite dimensional models.

The direct and maternal genetic correlation estimates between weights from birth to adulthood obtained for both B-spline models were very similar. Thus, the genetic and maternal correlation obtained with model BSQ5353 are shown in Table 2. The genetic correlation estimates decreased with increasing distance between weights. The direct genetic correlations obtained were higher than 0.5 for most ages. These results suggest that selection for weights at any age should alter the adult weight of females in the same direction. Thus, adult weight would be subject to marked increases as a response to selection for higher weights at younger ages, such as weaning and yearling, traits that have received greater emphasis in selection indices used for beef cattle in Brazil.

Using Gelbvieh weight records from birth to 365 days of age, Iwaisaki *et al.* (2005) reported similar direct genetic correlations estimated by multi-trait analysis and random regression on spline functions. However, the estimated correlations between all weights were lower than those obtained in this study. In Brazil, Baldi *et al.* (2010), using B-spline functions, also reported higher direct genetic correlations between adjacent weights and the estimated correlations between weights of females from birth to 8 years of age were generally similar to those observed in this study.

Using B-spline functions for weights from birth to 820 days of age, Meyer (2005b) observed small oscillations in direct genetic and direct permanent environmental correlations at ages in which knots were inserted. Iwaisaki *et al.* (2005) stated that RRMs on linear spline functions allow to estimate genetic parameter with smoother trajectories as spline functions fitted the data better at ages with lower number of records.

The estimates of maternal genetic correlations between weights from birth to 365 days of age were high, ranging from 0.61 to 0.84. Similarly, correlation estimates between weights after 365 days of age until 8 years of age also presented a moderate-to-high magnitude (Table 2). These results suggest that maternal effects at early ages are controlled by the same genes. Also using B-spline functions, Baldi *et al.* (2010) reported higher maternal genetic correlations, especially between weights obtained after weaning (0.99). On the other hand, Albuquerque and Meyer (2001), using RRM fitting Legendre polynomials, observed low maternal genetic correlations between weight at birth and weights from weaning to 550 days of age, but high correlations between the other standard ages (from 240 to 550 days).

For both B-spline models, the correlations of animal permanent environmental effects decreased with increasing time interval between weights, with lower correlations being observed between weight at birth and weights after 730 days of age (Table 3). High animal permanent environmental correlation estimates were observed between

Table 2 *Estimates of direct additive genetic (above the diagonal) and maternal (below the diagonal) correlations between weights from birth to adulthood obtained for model BSQ5353*

Weight	BW	W120	W240	W365	W540	W730	W1095	W1460	W1825	W2190	W2555	W2920
BW	–	0.78	0.70	0.69	0.56	0.56	0.43	0.48	0.43	0.38	0.32	0.34
W120	0.79	–	0.83	0.60	0.56	0.62	0.54	0.50	0.52	0.53	0.49	0.43
W240	0.72	0.73	–	0.82	0.62	0.60	0.61	0.57	0.55	0.52	0.45	0.47
W365	0.61	0.74	0.84	–	0.87	0.76	0.70	0.61	0.62	0.63	0.50	0.52
W540	0.50	0.54	0.66	0.78	–	0.89	0.74	0.60	0.65	0.60	0.64	0.63
W730	0.40	0.50	0.54	0.71	0.79	–	0.80	0.60	0.61	0.63	0.67	0.57
W1095	0.41	0.48	0.44	0.58	0.69	0.82	–	0.81	0.76	0.77	0.67	0.60
W1460	0.34	0.31	0.33	0.49	0.57	0.57	0.81	–	0.80	0.78	0.69	0.62
W1825	0.29	0.24	0.32	0.50	0.55	0.49	0.79	0.83	–	0.77	0.87	0.89
W2190	0.21	0.26	0.29	0.54	0.57	0.55	0.77	0.71	0.82	–	0.95	0.90
W2555	0.20	0.22	0.20	0.39	0.52	0.57	0.67	0.74	0.72	0.93	–	0.92
W2920	0.21	0.18	0.20	0.35	0.50	0.50	0.69	0.63	0.88	0.94	0.97	–

BW: weight at birth.
W120, W240, W365, W540, W730, W1095, W1460, W1825, W2190, W2555 and W2920: weight at 120, 240, 365, 540, 730, 1095, 1460, 1825, 2190, 2555 and 2920 days of age, respectively.

Table 3 *Estimates of animal environment (above the diagonal) and maternal permanent environment (below the diagonal) correlations between weights from birth to adult age obtained for model BSQ5353.*

Weight	BW	W120	W240	W365	W540	W730	W1095	W1460	W1825	W2190	W2555	W2920
BW	–	0.72	0.66	0.62	0.60	0.49	0.42	0.42	0.33	0.31	0.32	0.30
W120	0.67	–	0.75	0.72	0.70	0.55	0.44	0.47	0.49	0.42	0.41	0.40
W240	0.59	0.68	–	0.75	0.73	0.64	0.60	0.61	0.64	0.62	0.60	0.56
W365	0.55	0.64	0.62	–	0.70	0.69	0.67	0.70	0.62	0.60	0.64	0.63
W540	0.45	0.62	0.53	0.56	–	0.62	0.60	0.62	0.51	0.50	0.55	0.57
W730	0.48	0.50	0.48	0.62	0.49	–	0.79	0.76	0.68	0.64	0.60	0.61
W1095	0.35	0.42	0.46	0.54	0.44	0.68	–	0.74	0.78	0.64	0.65	0.67
W1460	0.34	0.34	0.43	0.43	0.51	0.66	0.59	–	0.81	0.79	0.76	0.80
W1825	0.33	0.45	0.49	0.41	0.50	0.57	0.56	0.60	–	0.83	0.84	0.87
W2190	0.32	0.31	0.46	0.49	0.47	0.44	0.52	0.67	0.64	–	0.87	0.88
W2555	0.30	0.32	0.33	0.44	0.42	0.42	0.48	0.58	0.65	0.66	–	0.92
W2920	0.28	0.27	0.32	0.40	0.39	0.38	0.39	0.67	0.69	0.75	0.74	–

BW = weight at birth.
W120, W240, W365, W540, W730, W1095, W1460, W1825, W2190, W2555 and W2920: weight at 120, 240, 365, 540, 730, 1095, 1460, 1825, 2190, 2555 and 2920 days of age, respectively.

weights from birth to 540 days of age and from 730 days of age until 8 years of age. Maternal permanent environmental correlation estimates between weight at birth and the other weights were of medium to high magnitude, and were higher between final ages. The maternal permanent environmental correlations estimated at the beginning of the study period differed from those reported by Baldi *et al.* (2010), who observed negative correlations between weights from birth to 12 months of age, but close to unity between weights after 18 months of age.

The three main eigenfunctions for direct additive genetic effect obtained with model BSQ5353 are shown in Figure 4. The three main eigenvalues of the coefficient matrix for direct additive genetic effects were responsible for 98.7% of the total direct additive genetic variation, with the first and second eigenvalue accounting for 59.2% of this variation. The first and second eigenfunctions for direct additive genetic effect were positive throughout the trajectory (Figure 4), indicating that selection based on these eigenfunctions will affect weights at all ages in the same direction. The third eigenfunction for direct additive genetic effect was negative from 60 to 300 days of age and positive during the remaining period. This sign change indicates that selection on this component will have opposite effects at the beginning and at the end of the trajectory. However, the possibility of genetic change because of selection on this eigenfunction is small as its eigenvalue was responsible for only 5.38% of additive genetic variance. Similar results have been reported by Baldi *et al.* (2010) for weights of Canchim females from birth to 8 years of age.

Using weights of Angus females after 365 days of age, Meyer (2005c) reported a wider fluctuation in the trajectory of the first two eigenfunctions for models with a larger number of segments or knots. According to Meyer (2005b), eigenfunctions for direct additive genetic and direct permanent environmental effects obtained with RRMs using B-spline functions were more consistent or regular than those obtained with ordinary polynomials because of the better fit of B-splines.

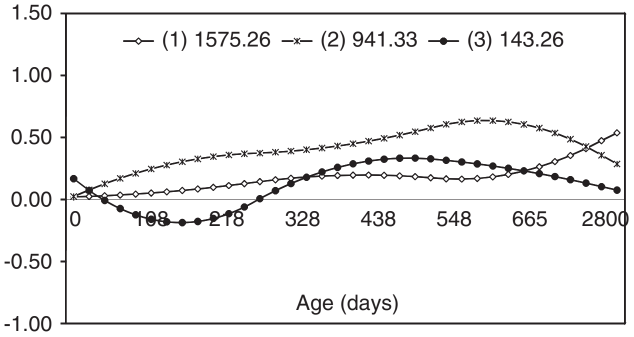

Figure 4 Eigenfunctions (1, 2 and 3) estimated for direct additive genetic and their respective eigenvalues obtained for model BSQ5353 (order of fit of five for direct genetic and permanent environmental effects of the animal and of three for maternal genetic and permanent environmental effects, with five classes of residual variances).

RRMs using quadratic B-spline functions, as base functions, fitted the data similarly to Legendre polynomials, but higher number of parameters needed to be estimated with the former. However, the convergence was reached faster with B-spline functions than Legendre polynomials. According to Meyer (2005b) and Misztal (2006), applying spline functions the computational requirements are less than for corresponding analyses with the same number of coefficients, but using Legendre polynomial, because the matrices in the mixed model are sparser with spline functions.

In the data set studied, a model fitting quadratic B-splines, with four knots or three segments for direct genetic and animal permanent environmental effects and two knots for maternal genetic and permanent environmental effects, was a parsimonious model and was able to describe the covariance structure adequately.

Conclusions

A model fitting quadratic B-splines, with four knots or three segments for direct genetic and animal permanent environmental effects and two knots for maternal genetic and maternal permanent environmental effects, was the most adequate and parsimonious model to describe the covariance structure of the data. Despite the estimates of variance components and genetic parameter obtained by RRMs on Legendre polynomials, B-spline functions and multi-trait models were similar; RRMs on B-spline functions required less computational requirement and were time consuming. These aspects are very important to define an appropriate model for genetic evaluation in large scale.

Selection for higher weight, such as at young ages, should be performed taking into account an increase in mature cow weight. Particularly, this is important in most of Nellore beef cattle production systems, where the cow herd is maintained on range conditions. The likelihood to modify the growth curve in Nellore cattle in order to obtain animals with fast growth at young ages and moderate-to-low mature cow weight is restricted.

Acknowledgments

The authors thank the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). We are indebted to the Associação Nacional de Criadores e Pesquisadores (ANCP) for the use of the data set.

References

Albuquerque LG and Meyer K 2001. Estimates of covariance functions for growth from birth to 630 days of age in Nellore cattle. Journal of Animal Science 79, 2776–2789.

Arango JA, Cundiff LV and Van Vleck LD 2004. Covariance functions and random regression models for cow weight in beef cattle. Journal of Animal Science 82, 54–67.

Baldi FS, Alencar MM and Albuquerque LG 2010. Random regression analyses using B-splines functions to model growth from birth to adult age in Canchim cattle. Journal of Animal Breeding and Genetics 127, 433–441.

Boligon AA, Albuquerque LG, Mercadante MEZ and Lobo RB 2009. Herdabilidades e correlações entre pesos do nascimento a idade adulta em rebanhos da raça Nelore. Revista Brasileira de Zootecnia 38, 2320–2326.

Boligon AA, Mercadante MEZ, Forni S, Lobo RB and Albuquerque LG 2010. Covariance functions for weights from birth to maturity in Nellore cows. Journal of Animal Science 88, 849–859.

De Boor C 2001. A practical guide to splines, vol. 27 of Springer Series in Applied Mathematics, 2nd edition. Springer Verlag, New York.

Iwaisaki H, Tsuruta S, Misztal I and Bertrand JK 2005. Genetic parameters estimated with multi-trait and linear splines-random regression models using Gelbvieh early growth data. Journal of Animal Science 83, 499–506.

Jenkins TG and Ferrell CL 1994. Productivity through weaning of nine breeds of cattle under varying feed availabilities: I. Initial evaluation. Journal of Animal Science 72, 2787–2797.

Marx BD and Eilers PHC 1998. Direct generalized additive modelling with penalized likelihood. Computational Statistic and Data Analysis 28, 193–209.

Meyer K 1999. Estimates of genetic and phenotypic covariance functions for post weaning growth and mature weight of beef cows. Journal of Animal Science 116, 181–205.

Meyer K 2000. Random regressions to model phenotypic variation in monthly weights of Australian beef cows. Livestock Production Science 65, 19–38.

Meyer K 2001. Estimates of direct and maternal covariance functions for growth of Australian beef calves from birth to weaning. Genetic Selection Evolution 33, 487–514.

Meyer K 2004. Scope for a random regression model in genetic evaluation of beef cattle for growth. Livestock Production Science 86, 69–83.

Meyer K 2005a. Estimates of genetic covariance functions for growth of Angus cattle. Journal of Animal Breeding and Genetics 122, 73–85.

Meyer K 2005b. Random regression analyses using B-splines to model growth of Australian Angus cattle. Genetic Select Evolution 37, 473–500.

Meyer K 2005c. Advances in methodology for random regression analyses. Australian Journal of Experimental Agriculture 45, 847–859.

Meyer K 2006. 'WOMBAT' – digging deep for quantitative genetic analyses by restricted maximum likelihood. Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brazil, CD-ROM.

Misztal I 2006. Properties of random regression models using linear splines. Journal of Animal Breeding and Genetics 123, 74–80.

Mrode RA 2005. Linear models for the prediction of animal breeding values, 2nd edition, CABI Publishing, Cambridge, MA, USA, 344pp.

Nephawe KA 2004. Application of random regression models to the genetic evaluation of cow weight in Bonsmara cattle of South Africa. South African Journal of Animal Science 34, 166–173.

Nobre PRC, Misztal I, Tsuruta S, Bertrand JK, Silva LOC and Lopes PS 2003. Analyses of growth curves of Nellore cattle by multiple-trait and random regression models. Journal of Animal Science 81, 918–926.

Rice JA and Wu CO 2001. Nonparametric mixed effects models for unequally sampled noisy curves. Biometrics 57, 253–259.

Ruppert D, Wand MP and Carroll RJ 2003. Semiparametric regression. Cambridge University Press, New York.

Sakaguti ES, Silva MA, Quaas RL, Martins EN, Lopes PS and Silva LOC 2003. Avaliação do crescimento de bovinos jovens da raça Tabapuã, por meio de análises de funções de covariância. Revista Brasileira de Zootecnia 32, 864–874.

Sánchez JP, Misztal I, Aguilar I and Bertrand JK 2008. Genetic evaluation of growth in a multibreed beef cattle population using random regression-linear splines models. Journal of Animal Science 86, 267–277.

Schwarz G 1978. Estimating the dimension of a model. Annals of Statistics 6, 461–464.

Tier B and Meyer K 2004. Approximating prediction error covariances in multiple-trait and random regression models. Journal of Animal Breeding and Genetics 121, 77–89.