

# Polynomials to model the growth of young bulls in performance tests

D. C. B. Scalez<sup>1,a</sup>, B. O. Fragomeni<sup>2</sup>, T. L. Passafaro<sup>2</sup>, I. G. Pereira<sup>2</sup> and F. L. B. Toral<sup>2†</sup>

<sup>1</sup>Faculdade de Agronomia, Medicina Veterinária e Zootecnia, Universidade Federal de Mato Grosso, Avenida Fernando Correa da Costa, 2367, CEP 78060-900, Cuiabá/MT, Brazil; <sup>2</sup>Departamento de Zootecnia, Escola de Veterinária, Universidade Federal de Minas Gerais (DZOO/EV/UFGM), Avenida Antônio Carlos, 6627, Caixa Postal 567, CEP 31270-901, Belo Horizonte/MG, Brazil.

(Received 7 August 2013; Accepted 25 September 2013; First published online 10 January 2014)

---

*The use of polynomial functions to describe the average growth trajectory and covariance functions of Nellore and MA (21/32 Charolais + 11/32 Nellore) young bulls in performance tests was studied. The average growth trajectories and additive genetic and permanent environmental covariance functions were fit with Legendre (linear through quintic) and quadratic B-spline (with two to four intervals) polynomials. In general, the Legendre and quadratic B-spline models that included more covariance parameters provided a better fit with the data. When comparing models with the same number of parameters, the quadratic B-spline provided a better fit than the Legendre polynomials. The quadratic B-spline with four intervals provided the best fit for the Nellore and MA groups. The fitting of random regression models with different types of polynomials (Legendre polynomials or B-spline) affected neither the genetic parameters estimates nor the ranking of the Nellore young bulls. However, fitting different type of polynomials affected the genetic parameters estimates and the ranking of the MA young bulls. Parsimonious Legendre or quadratic B-spline models could be used for genetic evaluation of body weight of Nellore young bulls in performance tests, whereas these parsimonious models were less efficient for animals of the MA genetic group owing to limited data at the extreme ages.*

---

**Keywords:** B-splines, beef cattle, legendre, random regression

## Implications

Adjustments of the average growth trajectories and covariance functions are useful tools for longitudinal traits, such as cattle body weight. The average growth trajectories and covariance functions for the body weight of young bulls must be carefully adjusted in performance tests before genetic evaluation because the polynomial type and the number of parameters might affect the genetic parameter estimates and ranking of the animals.

## Introduction

The main information available for beef cattle genetic evaluation and selection in Brazil are body weights measured over time. More recently, estimates of functions that describe animal growth have become more important because of the increased interest in using random regression models for the genetic evaluation of traits measured repeatedly over the

animal's life (Meyer, 2005). According to Kirkpatrick *et al.* (1990), the average growth trajectory can be defined as the mean of the body weight of the population as a function of age. The association of body weight with age can be assessed using random regression models.

Polynomials have frequently been used in studies conducted using beef cattle growth data, and the general functions that describe the average growth trajectory and the direct additive genetic and permanent environmental effects have been modelled using orthogonal polynomials, such as Legendre polynomials (Bohmanova *et al.*, 2005; Boligon *et al.*, 2010; Baldi *et al.*, 2010a), or by B-spline polynomials (Meyer, 2005; Baldi *et al.*, 2010b; Boligon *et al.*, 2012).

Most of the studies have focused only on adjustment of the covariance functions and ignored adjustment of the average growth trajectory. Once the covariance functions depend on deviations in the growth trajectory of each animal relative to the average growth trajectory of its group (Schaeffer and Dekkers, 1994), it is possible that a model with a poorly fit average growth trajectory results in poorly fit covariance functions, regardless of the number of parameters considered. The study presented here used polynomial functions to describe the fixed average growth

---

<sup>a</sup> Present address: Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista 'Julio de Mesquita Filho' – Campus de Jaboticabal, Via de acesso Prof. Paulo Donato Castellane, CEP 14884-900, Jaboticabal/SP

<sup>†</sup> E-mail: flbtoral@ufmg.br

trajectory and the random additive genetic and permanent environmental effects for Nellore and MA (21/32 Charolais + 11/32 Nellore) Brazilian young bulls subjected to performance tests to best fit the random regression models.

## Material and methods

The weight and age data of two genetic groups of young bulls (Nellore and MA) recorded during performance tests were used. The genetic group MA (21/32 Charolais + 11/32 Nellore, approximately) is a composite group from a rotational crossbreeding of Canchim (5/8 Charolais + 3/8 Zebu) and Charolais bulls with Nellore cows. The Nellore breed database consisted of 16 291 observations of 3356 animals from 37 performance tests conducted by the Grupo Proवादos a Pasto. Of the total, 15% and 85% of the animals present in the data set had four and five records each, respectively. The performance tests were conducted on ranches in the state of Goiás, Brazil, from 1997 through 2009 according to the official norms of the Brazilian Association of Zebu Breeders (Josahkian *et al.*, 2009). The weaned animals were kept in a single lot on pasture with mineral supplementation throughout the evaluation period. They were weighed after an adaptation period of 70 days and at the end of the test (224 days after the end of the adaptation period). They were also weighed at two or three intermediate times.

The data for the MA young bulls came from the evaluation programme of young bulls from the Agropecuária Ipameri. The young bulls' performance was assessed through 10 performance tests conducted between weaning (225 days of age, on average) and ~18 months of age. The tests were conducted from 1997 through 2007, except for 1999, on the Santa Helena Ranch in Jussara, Goiás, Brazil. The animals were kept on pasture and received protein-energetic supplementation during the dry season and mineral supplementation throughout the evaluation period. The animals were weighed at the beginning (weaning), middle (at least two weighings) and end of the test. The database consisted of 3997 observations of 884 MA young bulls. Of the total, 3%, 63%, 13% and 21% of the animals present in the data set had three, four, five and six records each, respectively.

The two databases were used to fit the average growth trajectories using linear, quadratic, cubic, quartic and quintic Legendre polynomials, and also quadratic B-splines with two, three and four regular intervals. The polynomials were always nested in each performance test because the minimum and maximum ages were not constant and genotype–environment interaction may change the shape of the average growth trajectory. In addition to the linear through quintic covariables, the Legendre polynomials also included an intercept term. Legendre coefficients were obtained using the Abramowitz and Stegun (1965) equation. Thus, those models that used Legendre polynomials were named according to the number of covariables used as LEG2, LEG3, LEG4, LEG5 and LEG6 for the linear through quintic polynomials, respectively.

The coefficients of the quadratic B-spline functions were generated for the  $k$  interval defined by the points  $T_k$  and  $T_{k+1}$ , with  $T_k \leq T_{k+1}$ , according to De Boor (2001) and Meyer (2005). In the present study, the polynomial functions considered were those with up to four equally spaced intervals. The division into  $m$  intervals requires the specification of  $m-1$  internal knots and two external knots ( $T_0$  and  $T_m$ ). This approach generates  $m+1$  knots and  $m+p$  non-null functions  $\phi_{k,p}$  (Meyer, 2005). The quadratic B-spline models were named according to the number of non-null functions. Thus, the quadratic B-splines with two, three and four regular intervals were named QBS4, QBS5 and QBS6, respectively.

### Evaluation of fixed-effect model

First, the regression coefficients for each model were estimated using the restricted maximum likelihood method (REML) without additive genetic and permanent environmental effects. After estimating the regression coefficients and the weights at each age, the residues were calculated for each observation to obtain the fit criteria. The following criteria for comparing the models were considered: the coefficient of determination ( $R^2$ ) obtained from the squared correlation among the observed and estimated values; the mean absolute deviation ( $MAD = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$ ), where  $n$  represents the number of observations and  $y_i$  and  $\hat{y}_i$  represent the observed and estimated values, respectively; the mean square error ( $MSE = \frac{\sum_{i=1}^n \hat{\epsilon}_i^2}{n}$ ), where  $\hat{\epsilon}_i$  represents the errors estimated for each observation; the value of the restricted likelihood function ( $-2RLL = (n-p) \log(2\pi) + \log|V| + \log|X'V^{-1}X| + (y - X\tilde{\beta})' V^{-1} (y - X\tilde{\beta})$ ), where  $p$  represents the rank of the fixed-effect matrix  $X$ ,  $V$  is the matrix of (co)variance of  $y$  (which represents the vector with the observations),  $\tilde{\beta}$  is the fixed-effect vector and  $k$  is the number of random parameters of the model; the Akaike information criterion ( $AIC = -2RLL + 2(p+k)$ ); and the consistent Akaike information criterion ( $CAIC = -2RLL + (p+k)[\text{Log}(n) + 1]$ ).

### Fixed and random regression models

In a second set of analyses, some alternatives to random regression models for modelling the average growth trajectory and covariance functions using Legendre and quadratic B-spline polynomials were evaluated. The general random regression model with Legendre polynomials used was as follows:

$$y_{ijm} = CG_i + \sum_{l=0}^{K_{agt}} \phi_l(\text{age}_j)_{(i^*)} \beta_{l(i^*)} + \sum_{l=0}^{K_a} \phi_l(\text{age}_j) \alpha_{lm} + \sum_{l=0}^{K_p} \phi_l(\text{age}_j) \rho_{lm} + e_{ijm} \quad (1)$$

where  $y_{ijm}$  represents the weight record of animal  $m$  at age  $j$  and in the contemporary group  $i$ ;  $CG_i$  is the classificatory fixed effect of contemporary group  $i$  (a contemporary group included

animals born in the same year-season and weighed on the same day);  $\phi_l(\text{age}_j)$  is the  $l$ th covariate at age  $j$ , which was nested in fixed-effect level  $i^*$  only to model the average growth trajectory ( $i^*$  represents the performance test and is part of the contemporary group definition);  $\beta_{l(i^*)}$  is the  $l$ th coefficient of fixed regression nested in fixed effect  $i^*$ ;  $\alpha_{lm}$  is the  $l$ th random regression coefficient for additive genetic effect of animal  $m$ ;  $\rho_{lm}$  is the  $l$ th random regression coefficient for permanent environmental effect of animal  $m$ ;  $e_{ijm}$  is a random error; and  $K_{agt}$ ,  $K_a$  and  $K_p$  represent orders of Legendre polynomials for the model fixed average growth trajectory and random additive genetic and permanent environmental effects, respectively. Random regression models with  $K_{agt}=K_a=K_p$ , of order from cubic to quintic were evaluated (LEG4, LEG5 and LEG6).

The general random regression model with quadratic B-splines was as follows:

$$y_{ijm} = CG_i + \sum_{k=0}^{M_{agt}+2} \phi_{k,2}(\text{age}_j)_{(i^*)} \beta_{k(i^*)} + \sum_{k=0}^{M_a+2} \phi_{k,2}(\text{age}_j) \alpha_{km} + \sum_{k=0}^{M_p+2} \phi_{k,2}(\text{age}_j) \rho_{km} + e_{ijm} \quad (2)$$

where  $\phi_{k,2}(\text{age}_j)$  represents the  $k$ th quadratic B-spline covariate at age  $j$ , which was nested in fixed-effect level  $i^*$  only to model the average growth trajectory;  $\beta_{k(i^*)}$  is  $k$ th coefficient of fixed regression nested in fixed effect  $i^*$ ;  $\alpha_{km}$  is the  $k$ th random regression coefficient for the additive genetic effect of animal  $m$ ;  $\rho_{km}$  is the  $k$ th random regression coefficient for the permanent environmental effect of animal  $m$ ; and  $M_{agt}$ ,  $M_a$  and  $M_p$  represent numbers of regular intervals used to model the fixed average growth trajectories and random additive genetic and permanent environmental effects, respectively. Random regression models with  $M_{agt}=M_a=M_p$  from two to four intervals were evaluated (QBS4, QBS5 and QBS6).

In matrix notation, the Legendre (1) and quadratic B-spline (2) general random regression models were as follows:

$$\underset{\sim}{y} = \underset{\sim}{X} \underset{\sim}{\beta} + \underset{\sim}{Z_1} \underset{\sim}{\alpha} + \underset{\sim}{Z_2} \underset{\sim}{\rho} + \underset{\sim}{e} \quad (3)$$

where  $\underset{\sim}{y}$  represents the vector of observations;  $\underset{\sim}{X}$  is the incidence matrix of fixed effects (contemporary groups and average growth trajectories),  $\underset{\sim}{\beta}$  is the vector of fixed-effect solutions;  $\underset{\sim}{Z_1}$  is the incidence matrix of random additive genetic effects;  $\underset{\sim}{\alpha}$  is the vector of random additive genetic effects solutions;  $\underset{\sim}{Z_2}$  is the incidence matrix of random permanent environmental effects;  $\underset{\sim}{\rho}$  is the vector of random permanent environmental effects solutions; and  $\underset{\sim}{e}$  is the vector of random errors.

The covariance of random effects was defined as follows:

$$V \begin{bmatrix} \underset{\sim}{\alpha} \\ \underset{\sim}{\rho} \\ \underset{\sim}{e} \end{bmatrix} = \begin{bmatrix} K_\alpha \otimes A & 0 & 0 \\ 0 & K_p \otimes I & 0 \\ 0 & 0 & R \end{bmatrix} \quad (4)$$

where  $K_\alpha$  represents the covariance matrix of random regression coefficients for additive genetic effects;  $A$  is the additive genetic relationship matrix;  $K_p$  is the covariance matrix of random regression coefficients for permanent environmental effects;  $I$  is the identity matrix; and  $R$  is the covariance matrix of random error.

The numerator relationship matrix was constructed from pedigree data that consisted of an animal's data and data for some of its ancestors. The ancestors retained in the pedigree were those that were parents of the animals with data or that were connected to other animals in the pedigree. Thus, the pedigree was composed of only informative animals and these were 4283 animals for the Nellore group and 1491 for the MA group.

Residual variances were modelled considering four age classes (262 to 357, 358 to 452, 453 to 547 and 548 to 642 days for Nellore and 166 to 301, 302 to 437, 438 to 573 and 574 to 706 days for MA). The covariance components and genetic parameters were estimated by the REML method using the Wombat software (Meyer, 2006). The criteria -2RLL, AIC and CAIC were used to compare the models.

The young bulls' expected progeny difference for weights at 450 and 550 days were predicted using different models (Legendre or quadratic B-splines) to simulate a selection process in each performance test. The top 10% (TOP10) of each performance test were selected according to each model, and the percentage of concordance in the TOP10 of the two different models was calculated.

## Results and discussion

It was observed that body weights increased with age in Nellore and MA young bulls and the distribution of records over ages varied between the two genetic groups (Figures 1 and 2). Approximately 85% and 56% of the records were between 300 and 550 days in Nellore and MA, respectively. This interval contains target ages considered for post-weaning selection (365, 450 and 550 days) in Brazil and

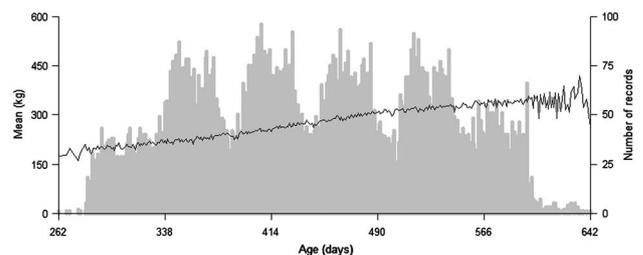


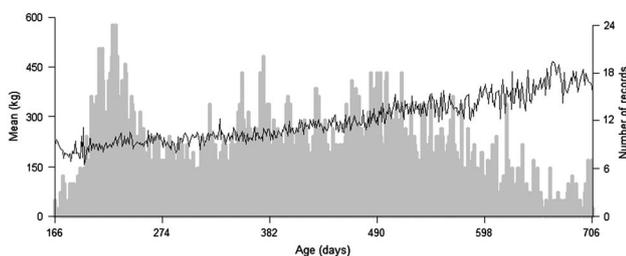
Figure 1 Number of records (bars) and mean body weight (black line) according to age for Nellore young bulls in performance tests.

represents 66% and 46% of the total intervals in Nellore and MA young bulls, respectively.

#### Evaluation of fixed-effect model

According to the  $R^2$ , MAD and MSE criteria, for a given type of polynomial model, the models that used more parameters were the best for the two groups (Table 1). These criteria did not penalise the models for the number of parameters; thus, they favoured the more complex models. When different types of polynomials using the same number of parameters were compared, no difference was detected for the  $R^2$  and MAD criteria. However, in comparisons that considered the MSE, the QBS6 model, which presented the same number of parameters as LEG6, was the best to fit for the Nellore breed. Considering the MA genetic group, the MSE criteria exhibited the lowest values for LEG6.

The Legendre polynomials provided a better fit for the data than the quadratic B-splines only when higher orders were used. In the case of the Legendre polynomials, for which the most suitable model was also LEG6, higher orders may have been necessary because of the unsuitability of the model for modelling the cyclical pattern in the data that is characteristic of growth traits that are subject to seasonal and climatic effects (Arango *et al.*, 2004). According to Kirkpatrick *et al.* (1994) and Meyer (1998), Legendre polynomials of a



**Figure 2** Number of records (bars) and mean body weight (black line) according to age for MA young bulls in performance tests.

high degree must be used to model growth curves, and thus a large number of parameters should be estimated. High-degree polynomials are also associated with numerical problems and have highly demanding computational requirements. The computational requirements can be reduced using B-spline functions, which can improve convergence and reduce computing costs (Bohmanova *et al.*, 2005). Regardless of the type of polynomial, the models that use more parameters exhibited the best fit for the two genetic groups according to the  $R^2$ , MAD and MSE criteria, whereas LEG2 presented the poorest fit (Table 1).

For the same type of polynomial, the  $-2RLL$  and AIC criteria demonstrated that LEG6 and QBS6 provided the best fit for the Nellore data. LEG5 and LEG6 fit the data equally well, and QBS6 provided the best fit of the quadratic B-splines for the MA data (Table 2). The CAIC, among the criteria used, imposed the greatest penalty with an increasing number of parameters (Bozdogan, 1987 and Wolfinger, 1993). Comparing the same type of polynomial, this criterion identified LEG4 and LEG5 as the most suitable models for the Nellore and MA groups, respectively. With respect to the comparisons of the quadratic B-splines polynomials, QBS5 exhibited the best fit for both genetic groups (Table 2).

Considering the  $-2RLL$ , AIC and CAIC criteria for general comparisons involving all of the models assessed, the quadratic B-spline polynomials were the best-fitting polynomials.  $-2RLL$  and AIC identified QBS6 as the model that best fit the growth data, and CAIC indicated that QBS5 and QBS6 fit the data for the two genetic groups (Table 2). A detailed analysis of the average growth trajectories indicated that the weights estimated by LEG6 were lower until 291 days of age (1.8 kg on average) and greater after an age of 615 days (3.2 kg on average) compared with the weights estimated by QBS6 for the Nellore breed. During the interval from 292 until 614 days, the mean of the differences among the weights estimated by the two models was nearly zero.

**Table 1** NRC, coefficient of determination ( $R^2$ ), MAD (kg) and MSE ( $kg^2$ ) according to the type and order of polynomial to model fixed average growth trajectories of Nellore and MA young bulls in performance tests

| Model <sup>a</sup>  | Nellore |       |      |       | MA  |       |      |        |
|---------------------|---------|-------|------|-------|-----|-------|------|--------|
|                     | NRC     | $R^2$ | MAD  | MSE   | NRC | $R^2$ | MAD  | MSE    |
| Legendre            |         |       |      |       |     |       |      |        |
| LEG2                | 74      | 0.723 | 23.2 | 857.8 | 20  | 0.730 | 27.6 | 1184.6 |
| LEG3                | 111     | 0.741 | 22.3 | 801.9 | 30  | 0.780 | 24.4 | 966.1  |
| LEG4                | 148     | 0.747 | 22.0 | 785.0 | 40  | 0.784 | 24.1 | 948.3  |
| LEG5                | 185     | 0.750 | 21.9 | 773.5 | 50  | 0.791 | 23.7 | 916.3  |
| LEG6                | 222     | 0.752 | 21.7 | 766.5 | 60  | 0.792 | 23.6 | 910.5  |
| Quadratic B-splines |         |       |      |       |     |       |      |        |
| QBS4                | 148     | 0.748 | 22.0 | 782.0 | 40  | 0.784 | 24.1 | 944.9  |
| QBS5                | 185     | 0.750 | 21.9 | 775.5 | 50  | 0.790 | 23.8 | 920.6  |
| QBS6                | 222     | 0.753 | 21.7 | 765.5 | 60  | 0.792 | 23.6 | 911.5  |

NRC = number of regression coefficients; MAD = mean absolute deviation; MSE = mean square error.

<sup>a</sup>LEGX represents Legendre polynomials with order X-1 and QBSY represents quadratic B-splines with Y-2 intervals.

**Table 2** NRC, values of the restricted likelihood function ( $-2RLL$ ), AIC and CAIC according to the type and order of polynomial to model fixed average growth trajectories of Nellore and MA young bulls in performance tests

| Model <sup>a</sup>  | Nellore |                          |                          |                           | MA  |                          |                          |                           |
|---------------------|---------|--------------------------|--------------------------|---------------------------|-----|--------------------------|--------------------------|---------------------------|
|                     | NRC     | $-2RLL (\times 10^{-3})$ | AIC ( $\times 10^{-3}$ ) | CAIC ( $\times 10^{-3}$ ) | NRC | $-2RLL (\times 10^{-3})$ | AIC ( $\times 10^{-3}$ ) | CAIC ( $\times 10^{-3}$ ) |
| Legendre            |         |                          |                          |                           |     |                          |                          |                           |
| LEG2                | 74      | 156.0                    | 156.2                    | 156.8                     | 20  | 39.6                     | 39.6                     | 39.8                      |
| LEG3                | 111     | 154.8                    | 155.0                    | 156.0                     | 30  | 38.7                     | 38.8                     | 39.0                      |
| LEG4                | 148     | 154.3                    | 154.6                    | 155.9                     | 40  | 38.6                     | 38.7                     | 39.0                      |
| LEG5                | 185     | 154.0                    | 154.3                    | 155.9                     | 50  | 38.4                     | 38.5                     | 38.9                      |
| LEG6                | 222     | 153.7                    | 154.1                    | 156.1                     | 60  | 38.4                     | 38.5                     | 38.9                      |
| Quadratic B-splines |         |                          |                          |                           |     |                          |                          |                           |
| QBS4                | 148     | 153.9                    | 154.2                    | 155.6                     | 40  | 38.5                     | 38.6                     | 38.9                      |
| QBS5                | 185     | 153.6                    | 153.9                    | 155.5                     | 50  | 38.3                     | 38.4                     | 38.8                      |
| QBS6                | 222     | 153.1                    | 153.6                    | 155.5                     | 60  | 38.2                     | 38.3                     | 38.8                      |

NRC = number of regression coefficients; AIC = Akaike information criteria; CAIC = consistent Akaike information criterion.

<sup>a</sup>LEGX represents Legendre polynomials with order X-1 and QBSY represents quadratic B-splines with Y-2 intervals.

**Table 3** NCP, values of the restricted likelihood function ( $-2RLL$ ), AIC and CAIC according to the type and order of polynomial to model fixed average growth trajectories and random additive genetic and permanent environmental effects of Nellore and MA young bulls in performance tests

| Model <sup>a</sup>  | NCP | Nellore |        |        | MA      |        |        |
|---------------------|-----|---------|--------|--------|---------|--------|--------|
|                     |     | $-2RLL$ | AIC    | CAIC   | $-2RLL$ | AIC    | CAIC   |
| Legendre            |     |         |        |        |         |        |        |
| LEG4                | 24  | 98 043  | 98 091 | 98 300 | 25 777  | 25 825 | 26 000 |
| LEG5                | 34  | 97 915  | 97 984 | 98 279 | 25 704  | 25 772 | 26 020 |
| LEG6                | 46  | 97 782  | 97 874 | 98 274 | 25 658  | 25 750 | 26 086 |
| Quadratic B-splines |     |         |        |        |         |        |        |
| QBS4                | 24  | 97 599  | 97 647 | 98 855 | 25 652  | 25 700 | 25 875 |
| QBS5                | 34  | 97 376  | 97 444 | 97 739 | 25 554  | 25 622 | 25 870 |
| QBS6                | 46  | 97 180  | 97 272 | 97 671 | 25 491  | 25 583 | 25 918 |

NCP = number of covariance parameters; AIC = Akaike information criteria; CAIC = consistent Akaike information criterion.

<sup>a</sup>LEGX represents Legendre polynomials with order X-1 to model fixed average growth trajectory and random additive genetic and permanent environmental effects and QBSY represents quadratic B-splines with Y-2 intervals to model the same effects.

However, the weights estimated by LEG6 for the Nellore young bulls did not reflect the expected body weight at the final extremity of the age interval.

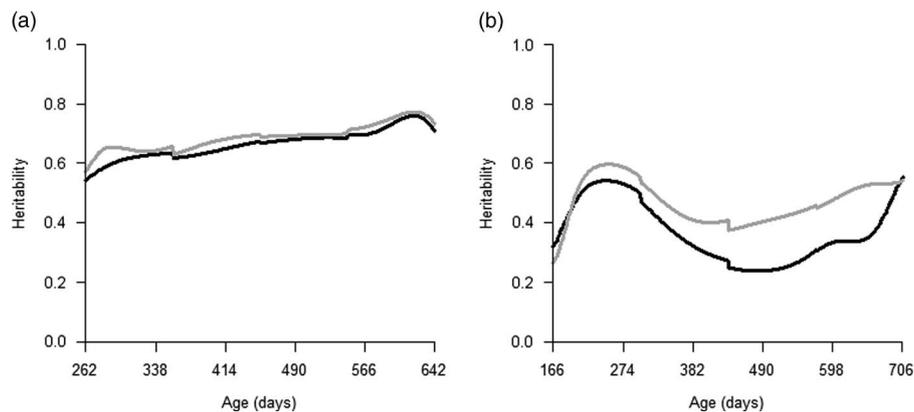
The weight predicted by LEG6 oscillated sharply at the final extremity of the age interval considered for the two groups (results not shown). Meyer (1999) and Arango *et al.* (2004) emphasised on the difficulties in the application of Legendre polynomials to cattle growth data, particularly because of the greater emphasis that they place on observations at the extremes of the growth curve. To overcome this problem, Meyer (2005) suggested using B-spline functions, which are curves that are constructed from lower-degree polynomial segments joined at points called knots. Furthermore, B-spline functions are capable of improving the adjustment at the extremes of the age interval because they fit each part of the trajectory according to the number of intervals established.

The QBS6 model described the average growth trajectory of the Nellore and MA young bulls more smoothly and consistently, particularly at the extremes of the age interval.

Therefore, this model was considered the most suitable because it exhibited a biological performance closer to that expected for the average growth trajectory (Figures 1 and 2). If Legendre polynomials were used, a better fit could be achieved by using higher-order polynomials. Using high-order Legendre polynomials (quintic, in this case) could have caused problems of biological interpretation of the curve shape, as discussed previously. There is no need to use B-splines with high-order coefficients, and there is better local control (Meyer, 2005) because a set of few functions is considered for each interval, whereas the Legendre polynomial fit one equation with many regression coefficients throughout the range of the independent variable (age).

#### Fixed and random regression models

Considering models with the same number of covariance parameters, performing random regression with quadratic B-splines to model fixed and random effects provided a better fit for the data compared with Legendre polynomials (Table 3). These results are in agreement with those



**Figure 3** Estimates of heritability for body weight of Nellore (a) and MA (b) young bulls in performance tests with quintic Legendre (grey line) and quadratic B-splines (black line) polynomials with four intervals to model fixed average growth trajectory and random additive genetic and permanent environmental effects.

presented by Meyer (2005) for the weight of Angus cattle from birth to 820 days. However, Baldi *et al.* (2010b) and Boligon *et al.* (2012) found that Legendre polynomials performed better than linear, quadratic and cubic B-splines to fit Canchim and Nellore weight data from birth to adult age. According to Meyer (2005), the lower degrees of B-splines in the individual segments and the local rather than global influence of individual observations partially explain the better fit of B-splines. The age intervals and distribution of records over the range of ages can also be important aspects of data fitting, and thus these factors may affect the type of polynomial that should be used. In addition to the wider range of ages, the distribution of weight records in the data set of Baldi *et al.* (2010b) and Boligon *et al.* (2012) was less stable than the distributions of records in the data sets of Meyer (2005) and the present data sets. These distributions might be an additional factor to consider before choosing the type of polynomial to be used in genetic evaluations.

The modelling of the fixed average growth trajectory and random additive genetic and permanent environmental effects by quadratic B-splines with three to four regular intervals (QBS5 and QBS6) provided a better fit with the weight data of Nellore and MA bulls submitted to performance tests than other models. Normally, a better fit can be achieved by increasing the number of intervals (Meyer, 2005; Baldi *et al.*, 2010b; Boligon *et al.*, 2012) because each quadratic B-spline function will be responsible for fitting a smaller interval. However, the minimum and mean individual numbers of the records may limit the slicing of an age interval. In random regression models, breeding values for growth are predicted over the age interval, but predictions for animals with few or poorly distributed records are strongly dependent on the animal's relatives' records. The minimum of three (MA) and four (Nellore) records and the mean of 4.5 (MA) and 4.8 (Nellore) records per animal might be sufficient to evaluate the current statistical models and to apply them for genetic evaluation of young bulls in performance tests.

The estimates of additive genetic and phenotypic variances for the growth of Nellore and MA young bulls increased over

the age interval. The estimates of permanent environmental and residual variances also increased over the age interval for the MA group, but they were constant for the Nellore animals. The additive genetic, permanent environmental, residual and phenotypic variances of weight in performance tests of Nellore and MA young bulls estimated by different Legendre models (LEG4, LE5 and LEG6) were very similar. Considering the Nellore data, the variance estimates were also similar for different quadratic B-spline models. The estimates of additive genetic and phenotypic variances obtained with the QBS6 model were lower than the estimates from the QBS4 and QBS5 models for MA young bulls (results not shown). The pattern of differences in variances were the same as reported for heritability below (Figure 3).

The estimates of the variances and heritabilities for the weight of Nellore young bulls in random regression models with quintic Legendre (LEG6) or quadratic B-spline (QBS6) polynomials were similar over the age interval (Figure 3a). The increase in heritability and the constant permanent environmental and residual variances indicate that the genetic differences among animals become more evident with age and that selection responses tend to be greater for weight at the end of the performance test for the Nellore group than for weight at other ages.

The estimates of heritability by random regression analysis of the Nellore weights obtained by Albuquerque and Meyer (2001) also tended to increase from weaning to 550 days. However, the estimates of permanent environmental and residual variances obtained by Albuquerque and Meyer (2001) were not constant. The young bulls in performance tests were kept in the same contemporary group over the total testing period, and changes in the general management were considered minimum. In general, commercial producers tend to change the type of management over time, such as by introducing nutritional supplementation. It is known that minimising changes in the environment is critical for estimating additive genetic variance and to select sires and dams for breeding.

The estimates of additive genetic variance for the weight of MA young bulls using LEG6 and QBS6 tended to increase

from the beginning of the age interval to 250 days and to decrease from 250 to 380 days of age. After 380 days of age, the additive variance increased with age, and additive genetic variances in the quadratic B-spline model (QBS6) were lower than the variances in the quintic Legendre (LEG6) model. Baldi *et al.* (2010b) found important maternal effects on the weight of Canchim beef cattle until 250 days of age; in the present study, these effects might have been confounded with direct additive genetic effects. However, it was not possible to model maternal effects in the present study because the ratio of young bulls to cow was 1.6 and there were no growth records for cows. A similar constraint was observed by Riley *et al.* (2007) with growth data of Brahman calves maintained in feedlot during 170 days after weaning. Riley *et al.* (2007) did not include maternal effects to model body weight of animals and reported inflated estimates of variances because of maternal effects.

Owing to changes in the variances, the estimates of heritabilities for the growth of MA young bulls in performance test (Figure 3b) rapidly increased from the beginning of the age interval to 250 days and decreased from 250 to 440 days of age. After this point, the estimates of heritabilities increased. The estimates of heritability obtained with the quintic Legendre (LEG6) model were markedly greater than the estimates from quadratic B-splines using four intervals (QBS6) after 380 days of age. At the end of the age intervals, the estimates of the heritability obtained with LEG6 and QBS6 became more similar because of the sudden increase in the estimates of heritabilities in the latter model. The estimates of genetic parameters for post-weaning body weight presented in this paper were obtained with data from 884 young bulls sired by 50 Charolais bulls. Of these Charolais bulls, 30 were unrelated or had missing parents. These bulls represented a broad sampling of bloodlines and included prominent bulls of the Charolais breed at that time. Riley *et al.* (2007) obtained estimates of genetic parameters with data on 724 steers and heifers from a progeny test on 27 Brahman bulls in the United States of America. Despite the small size of MA data (and Brahman data from Riley *et al.*, 2007), these results were obtained from a representative sample of MA young bull population and support selection for body weight in the whole population.

The genetic and phenotypic correlation estimates between weights from the beginning to the end of the performance test over the age interval were moderate to high, and they decreased with increasing distance between weights for the Nellore (Tables 4 and 5) and MA (Tables 6 and 7) bulls. Some unexpected oscillations in the additive genetic (Table 4) and phenotypic (Table 5) correlations among weights in the Nellore young bulls were observed. This situation was also noted by Meyer (2005) in the analysis of weight from birth to 820 days of age for Australian Angus cattle. According to Meyer (2005), some effects of joints between pieces of polynomials may have contributed to the oscillations. In general, the results of the present study corroborate those of Albuquerque and Meyer (2001) and Boligon *et al.* (2012), thereby suggesting that selection for body weight at any age

**Table 4** Estimates of genetic correlations between body weights of Nellore young bulls from the beginning to the end of age interval of performance tests with quintic Legendre polynomials (above diagonal) to model fixed average growth trajectory and random additive genetic and permanent environmental effects and quadratic B-splines with four regular intervals (below diagonal) to model the same effects

| Weight | W300  | W350  | W400  | W450  | W500  | W550  |
|--------|-------|-------|-------|-------|-------|-------|
| W300   | –     | 0.996 | 0.973 | 0.953 | 0.936 | 0.915 |
| W350   | 0.800 | –     | 0.988 | 0.973 | 0.959 | 0.945 |
| W400   | 0.375 | 0.831 | –     | 0.995 | 0.987 | 0.981 |
| W450   | 0.601 | 0.803 | 0.861 | –     | 0.998 | 0.992 |
| W500   | 0.759 | 0.538 | 0.353 | 0.772 | –     | 0.995 |
| W550   | 0.705 | 0.340 | 0.196 | 0.648 | 0.970 | –     |

W300, W350, W400, W450, W500 and W550: weight at 300, 350, 400, 450, 500 and 550 days of age, respectively.

**Table 5** Estimates of phenotypic correlations between body weights of Nellore young bulls from the beginning to the end of age interval of performance tests with quintic Legendre polynomials (above diagonal) to model fixed average growth trajectory and random additive genetic and permanent environmental effects and quadratic B-splines with four regular intervals (below diagonal) to model the same effects

| Weight | W300  | W350  | W400  | W450  | W500  | W550  |
|--------|-------|-------|-------|-------|-------|-------|
| W300   | –     | 0.923 | 0.865 | 0.838 | 0.810 | 0.775 |
| W350   | 0.900 | –     | 0.917 | 0.884 | 0.853 | 0.828 |
| W400   | 0.522 | 0.769 | –     | 0.914 | 0.884 | 0.856 |
| W450   | 0.357 | 0.624 | 0.928 | –     | 0.919 | 0.896 |
| W500   | 0.354 | 0.596 | 0.873 | 0.982 | –     | 0.928 |
| W550   | 0.511 | 0.686 | 0.904 | 0.928 | 0.925 | –     |

W300, W350, W400, W450, W500 and W550: weight at 300, 350, 400, 450, 500 and 550 days of age, respectively.

changes the weight at any other age in the same direction. The estimates of correlations among weights were lower in the quadratic B-spline model using four regular intervals than in the quintic Legendre polynomial model. To independently evaluate growth at different points of an animal's life, the process of obtaining covariables and modelling the fixed and random effects at different points of age intervals using B-splines should be more efficient than Legendre polynomials. One potential consequence is that breeding values for body weight obtained by quadratic B-spline random regression models might be an interesting option to avoid an indirect increase in adult weight being caused by selection for increasing weight at early ages.

In general, the genetic parameters estimates based on different models were similar, but it is important to evaluate the effects of these small differences on expected progeny differences and the selection process. Because the mean percentages of concordance in the TOP10 animals were greater than 93% in every comparison (Table 8), the number of covariance parameters and the type of polynomial likely did not affect the selection process for body weight in performance tests for the Nellore breed. The results

**Table 6** Estimates of genetic correlations between body weights of MA young bulls from the beginning to the end of age interval of performance tests with quintic Legendre polynomials (above diagonal) to model fixed average growth trajectory and random additive genetic and permanent environmental effects and quadratic B-splines with four regular intervals (below diagonal) to model the same effects

| Weight | W300  | W350  | W400  | W450  | W500  | W550  |
|--------|-------|-------|-------|-------|-------|-------|
| W300   | –     | 0.980 | 0.913 | 0.850 | 0.828 | 0.828 |
| W350   | 0.963 | –     | 0.976 | 0.937 | 0.920 | 0.915 |
| W400   | 0.892 | 0.974 | –     | 0.990 | 0.980 | 0.966 |
| W450   | 0.789 | 0.876 | 0.952 | –     | 0.996 | 0.980 |
| W500   | 0.691 | 0.741 | 0.830 | 0.953 | –     | 0.992 |
| W550   | 0.625 | 0.606 | 0.665 | 0.808 | 0.936 | –     |

W300, W350, W400, W450, W500 and W550: weight at 300, 350, 400, 450, 500 and 550 days of age, respectively.

**Table 7** Estimates of phenotypic correlations between body weights of MA young bulls from the beginning to the end of age interval of performance tests with quintic Legendre polynomials (above diagonal) to model fixed average growth trajectory and random additive genetic and permanent environmental effects and quadratic B-splines with four regular intervals (below diagonal) to model the same effects

| Weight | W300  | W350  | W400  | W450  | W500  | W550  |
|--------|-------|-------|-------|-------|-------|-------|
| W300   | –     | 0.820 | 0.773 | 0.711 | 0.689 | 0.663 |
| W350   | 0.798 | –     | 0.823 | 0.765 | 0.728 | 0.686 |
| W400   | 0.737 | 0.820 | –     | 0.806 | 0.775 | 0.732 |
| W450   | 0.655 | 0.723 | 0.778 | –     | 0.789 | 0.762 |
| W500   | 0.596 | 0.636 | 0.706 | 0.763 | –     | 0.816 |
| W550   | 0.535 | 0.548 | 0.615 | 0.694 | 0.768 | –     |

W300, W350, W400, W450, W500 and W550: weight at 300, 350, 400, 450, 500 and 550 days of age, respectively.

demonstrated the similarity of the estimates of variances and heritabilities based on models with different types of polynomials (Legendre or quadratic B-splines) or different numbers of covariance parameters. However, the differences in variances and heritability estimates for the weight of the MA young bulls in performance tests from different statistical models using different numbers of covariance parameters and types of polynomials appeared to affect the ranking and selection of young bulls (Table 8). The body weight records of the MA young bulls were distributed in a wider interval than the Nellore weights, and a sparse distribution in the last interval might have resulted in poor estimates of fixed and random effects in the MA database. Because of this sparsity, small differences in the adjustment of the fixed average growth trajectory can affect the genetic parameter estimates for the MA genetic group more than those for the Nellore group. This is a common situation that must be addressed in real data sets. In some cases, such constraints limit the model that can be used and its parameterisation.

The percentage of concordance indicates similarity among models and may be used as an additional parameter to compare models. For a real data set, this statistic does not

**Table 8** Means of percentage of concordance in TOP10 Nellore and MA young bulls in each performance test according to expected progeny difference for body weight at 450 (above diagonal) and 550 (below diagonal) days for each pair of statistical models<sup>a</sup>

|      | Nellore |      |      | MA   |      |       |
|------|---------|------|------|------|------|-------|
|      | LEG4    | LEG5 | LEG6 | LEG4 | LEG5 | LEG6  |
| LEG4 | –       | 96.8 | 94.9 | –    | 2.2  | 1.11  |
| LEG5 | 96.5    | –    | 95.4 | 0.0  | –    | 100.0 |
| LEG6 | 93.4    | 94.6 | –    | 0.0  | 88.1 | –     |
|      | QBS4    | QBS5 | QBS6 | QBS4 | QBS5 | QBS6  |
| QBS4 | –       | 95.1 | 95.7 | –    | 78.7 | 66.4  |
| QBS5 | 93.7    | –    | 95.7 | 87.3 | –    | 74.1  |
| QBS6 | 93.2    | 95.8 | –    | 77.3 | 71.2 | –     |
|      | LEG4    | LEG5 | LEG6 | LEG4 | LEG5 | LEG6  |
| QBS4 | 99.1    | –    | –    | 11.1 | –    | –     |
|      | 98.7    |      |      | 0.0  |      |       |
| QBS5 | –       | 98.5 | –    | –    | 81.0 | –     |
|      |         | 96.3 |      |      | 85.3 |       |
| QBS6 | –       | –    | 93.7 | –    | –    | 71.8  |
|      |         |      | 96.1 |      |      | 80.1  |

<sup>a</sup>LEGX represents a model with Legendre polynomials of order X-1 to model the fixed effect of average growth trajectory and random additive genetic and permanent environmental effects and QBSY represents a model with quadratic B-spline with Y-2 regular intervals to model the fixed effect of average growth trajectory and random additive genetic and permanent environmental effects.

indicate the best model. It only indicates putative differences in the selection process. Despite the difference among –2RLL, AIC and CAIC, genetic evaluation of the body weight of Nellore young bulls submitted to performance tests can be performed using both Legendre and quadratic B-splines and with the most parsimonious models because there is a high concordance in the top-ranked bulls. However, the percentage of concordance in the top-ranked MA bulls also indicated differences among the models. For this genetic group, and most likely all others with poor structures, the statistical evaluation criteria (–2RLL, AIC and CAIC) and biological interpretation should be considered. The results of the present paper suggest that random regression genetic evaluation of MA young bulls in performance tests should be performed with quadratic B-splines of three to four regular intervals.

The quadratic B-spline functions appear to facilitate estimation and simple biological interpretations (Baldi *et al.*, 2010b) and are less susceptible to problems that occur frequently when orthogonal polynomials are used. In the paper, using quadratic B-spline polynomials was satisfactory for the two genetic groups, both for performance tests with the greatest and least number of observations. Generally, greater data dispersion was observed for performance tests with smaller number of observations. However, the data for these performance tests should also be analysed, and more complex models could have difficulties fitting them because of the faults in the data structure. More robust models, such as

those with quadratic B-spline polynomials, provided a better fit for the data, even when the data structure was less consistent. Thus, the best fit obtained by the use of quadratic B-spline polynomials was justified by their flexibility in fitting seasonal growth fluctuations in the cattle (Meyer, 2000).

## Conclusions

The use of Legendre or quadratic B-spline polynomials was satisfactory for modelling the fixed average growth trajectory and random additive genetic and permanent environmental effects of the Nellore young bulls. It was also possible to choose a parsimonious model for genetic evaluation of the weight of the Nellore young bulls in performance tests. Differences regarding polynomial types and the number of covariance parameters affected the selection process for the MA genetic group, and the quadratic B-splines with three to four equally spaced intervals provided the best fit for the data.

## Acknowledgements

The authors thank Humberto de Freitas Tavares (Sucuri Ranch – Provados a Pasto), Maurício Mello de Alencar (Embrapa Cattle – Southeast) and João Paulo Porto (Santa Helena Ranch – Agropecuária Ipameri) for allowing the use of databases for this study. Pró-Reitoria de Pesquisa/Universidade Federal de Minas Gerais funded the translation of the paper, FAPEMIG funded facilities for this research (PPM-00456-11), CNPq funded Fabio Luiz Buranelo Toral (Process 502401/2009-5) and Breno de Oliveira Fragomeni (studentship) and CAPES funded Daiane Cristina Becker Scalez (studentship).

## References

Abramowitz M and Stegun IA 1965. Handbook of mathematical functions with formulas, graphs, and mathematical functions, 55th edition. National Bureau of Standards Applied Mathematics, Dover, NY, USA.

Albuquerque LG and Meyer K 2001. Estimates of covariance functions for growth from birth to 630 of agen 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 F, Albuquerque LG and Alencar MM 2010a. Random regression models on Legendre polynomials to estimate genetic parameters for weights from birth to adult age in Canchim cattle. *Journal of Animal Breeding and Genetics* 127, 289–299.

Baldi F, Alencar MM and Albuquerque LG 2010b. 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.

Bohmanova J, Misztal I and Bertrand JK 2005. Studies on multiple trait and random regression models for genetic evaluation of beef cattle for growth. *Journal of Animal Science* 83, 62–67.

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

Boligon AA, Mercadante MEZ, Lôbo RB, Baldi F and Albuquerque LG 2012. Random regression analyses using B-spline functions to model growth of Nellore cattle. *Animal* 6, 212–220.

Bozdogan H 1987. Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions. *Psychometrika* 52, 345–370.

De Boor C 2001. A practical guide to splines, 2nd edition. Springer Verlag, New York, NY, USA.

Josahkian LA, Lucas CH and Machado CHC 2009. Manual do serviço de registro genealógico das raças zebuínas e PMGZ. Associação Brasileira dos Criadores de Zebu, Uberaba, MG, Brazil.

Kirkpatrick M, Lofsvold D and Bulmer M 1990. Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics* 124, 979–993.

Kirkpatrick M, Hill WG and Thompson R 1994. Estimating the covariance structure of traits during growth and aging, illustrated with lactation in dairy cattle. *Genetics Research* 64, 57–69.

Meyer K 1998. Estimating covariance functions for longitudinal data using a random regression model. *Genetics Selection Evolution* 30, 221–240.

Meyer K 1999. Estimates of genetic and phenotypic covariance functions for post weaning growth and mature weight of beef cows. *Journal of Animal Breeding and Genetics* 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 2005. Random regression analyses using B-spline to model growth of Australian Angus cattle. *Genetics Selection Evolution* 37, 473–500.

Meyer K 2006. 'Wombat' – digging deep for quantitative genetics analysis by restricted maximum likelihood. In Book of abstracts of 8th World Congress on Genetics Applied to Livestock Production, 13–18 August 2006, Belo Horizonte, Brazil, pp. 248.

Riley DG, Coleman SW, Chase CC Jr, Olson TA and Hammond AC 2007. Genetic parameters for body weight, hip height, and the ratio of weight and hip height from random regression analyses of Brahman feedlot cattle. *Journal of Animal Science* 85, 42–52.

Schaeffer LR and Dekkers JCM 1994. Random regressions in animal models for test-day production in dairy cattle. In Proceedings of 5th World Congress on Genetics Applied to Livestock Production, 7–12 August 1994, Guelph, Canada, pp. 443–446.

Wolfinger RD 1993. Covariance structure selection in general mixed models. *Communications in Statistics – Simulation and Computation* 22, 1079–1106.