

JOURNAL OF ANIMAL SCIENCE

The Premier Journal and Leading Source of New Knowledge and Perspective in Animal Science

Analysis of beef cattle longitudinal data applying a nonlinear model

S. Forni, M. Piles, A. Blasco, L. Varona, H. N. Oliveira, R. B. Lôbo and L. G. Albuquerque

J Anim Sci 2007.85:3189-3197.

doi: 10.2527/jas.2006-677 originally published online Jul 20, 2007;

The online version of this article, along with updated information and services, is located on the World Wide Web at:

<http://jas.fass.org/cgi/content/full/85/12/3189>



American Society of Animal Science

www.asas.org

Analysis of beef cattle longitudinal data applying a nonlinear model¹

S. Forni,* M. Piles,† A. Blasco,‡ L. Varona,§ H. N. Oliveira,#
R. B. Lôbo,|| and L. G. Albuquerque*²

*Faculdade de Ciências Agrárias e Veterinárias, UNESP, Jaboticabal, SP, 14884900, Brazil;

†IRTA, Unidad de Cunicultura, Caldes de Montbui, 68140, Spain; ‡Departamento de Ciencia Animal, Universidad Politécnica de Valencia, 46071, Spain; §Centre UdL-IRTA, Lleida, 25198, Spain;

#Faculdade de Medicina Veterinária e Zootecnia, UNESP, Botucatu, SP, 18618000, Brazil; and

||Faculdade de Medicina, USP, Ribeirão Preto, SP, 14049900, Brazil

ABSTRACT: The objective of this work was to evaluate the Nelore beef cattle, growth curve parameters using the Von Bertalanffy function in a nested Bayesian procedure that allowed estimation of the joint posterior distribution of growth curve parameters, their (co)variance components, and the environmental and additive genetic components affecting them. A hierarchical model was applied; each individual had a growth trajectory described by the nonlinear function, and each parameter of this function was considered to be affected by genetic and environmental effects that were described by an animal model. Random samples of the posterior distributions were drawn using Gibbs sampling and Metropolis-Hastings algorithms. The data set consisted of a total of 145,961 BW recorded from 15,386 animals. Even though the curve parameters were estimated for animals with few records, given that the information from related animals and the structure of systematic effects were considered in the curve fitting, all mature BW predicted were suitable. A large additive

genetic variance for mature BW was observed. The parameter a of growth curves, which represents asymptotic adult BW, could be used as a selection criterion to control increases in adult BW when selecting for growth rate. The effect of maternal environment on growth was carried through to maturity and should be considered when evaluating adult BW. Other growth curve parameters showed small additive genetic and maternal effects. Mature BW and parameter k , related to the slope of the curve, presented a large, positive genetic correlation. The results indicated that selection for growth rate would increase adult BW without substantially changing the shape of the growth curve. Selection to change the slope of the growth curve without modifying adult BW would be inefficient because their genetic correlation is large. However, adult BW could be considered in a selection index with its corresponding economic weight to improve the overall efficiency of beef cattle production.

Key words: Bayesian analysis, beef cattle, growth curve, longitudinal data, selection, Von Bertalanffy

©2007 American Society of Animal Science. All rights reserved.

J. Anim. Sci. 2007. 85:3189–3197
doi:10.2527/jas.2006-677

INTRODUCTION

In the Brazilian beef cattle industry, animal growth has been evaluated using BW at standard ages, such as birth, weaning, and 540 d of age, or BW gains. Selection decisions are based on information from the beginning of the growth process. Body weights at different ages are highly and positively genetically correlated (Lôbo

et al., 2000; Albuquerque and Meyer, 2001), and an increase in mature BW can be expected, increasing the costs of maintaining a parent population. Selection for growth has a correlated response increasing adult BW in different species (Archer et al., 1998; Mignon-Gastreau et al., 2000; Piles et al., 2003). Including adult BW in selection programs would be valuable, but many animals reach slaughter weight before maturity so that individual adult BW can only be recorded for a few animals.

An alternative would be to use early predictors of adult BW estimated from individual growth curves. The usual functions applied to describe growth are based on differential equations that seek a biological interpretation (Arango and Van Vleck, 2002). The parameters of these equations have a biological meaning related to

¹This work was supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior.

²Corresponding author: lgalb@fcav.unesp.br

Received October 9, 2006.

Accepted July 18, 2007.

initial conditions, growth rate, or adult BW, all linked to economic aspects of production. Usually, the growth function is predicted for each animal and, as a second step, environmental effects and (co)variance components of the parameters are estimated, ignoring the adjustment errors from the first step. Information on relatives is not considered when the curve parameters are estimated (i.e., not all information is used, and animals with few records cannot be included in the analysis).

The objective of this work was to evaluate the Nelore cattle growth curve parameters using a nonlinear model in a nested Bayesian procedure that allowed estimation of the joint posterior distribution of growth curve parameters, their (co)variance components, and the environmental and additive genetic components affecting them.

MATERIALS AND METHODS

The data used in the present study were not obtained in experimental conditions. Animals whose data were collected were raised on different commercial farms, and non-special conditions or treatment were adopted in order to provide such records. Therefore, Animal Care and Use Committee approval was not required.

Data

Data were supplied by the ANCP (Associação Nacional de Criadores e Pesquisadores Ribeirão Preto, São Paulo, Brazil), which has been running the Nelore Breeding Program since 1987 (Lôbo et al., 2005). Calves were born throughout the year, with a natural concentration in spring and summer, and were weaned and weighed at approximately 210 d of age. Besides weaning, animals were weighed at 90-d intervals from birth to 540 d of age, and those selected for reproduction were continuously weighed every 90 d.

Records were extracted for animals that were reared on pasture without supplemental feeding, did not have a foster dam, did not receive veterinary treatment, had dams older than 2 and younger than 16 yr of age, had birth weights greater than 20 kg, had all BW within the range given by the mean of all animals with the same age \pm 3 SD, had at least 5 valid BW (i.e., records meeting the above conditions), and belonged to a contemporary group with at least 6 animals. The contemporary group included sex, herd, management group, and year and month of birth. The data set consisted of a total of 145,961 BW recorded from 15,386 animals, which were the offspring of 501 sires and 7,574 dams. Pedigree information was obtained for 5 generations, resulting in a total of 29,897 animals in the relationship matrix. A greater number of records were taken at the beginning of the growth process, until 540 d of age for males and females, and adult BW were available mainly from females. The data set had 5,661 BW records of 1,420 animals that were more than 4 yr of age.

The Model

A hierarchical model was applied to describe the growth curve of each animal. Each individual had a growth trajectory described by a nonlinear function, and each parameter of this function was considered to be affected by genetic and environmental effects that were described by a linear model. In the first stage of the model, it was assumed that individual growth curves followed the Von Bertalanffy function (Von Bertalanffy, 1957):

$$y_{ij} = a_i [1 - b_i \exp(-k_i t_j)]^3 + \varepsilon_{ij},$$

where y_{ij} is the observed BW of individual i at age j , a_i can be interpreted as the average mature BW maintained independent of short-term fluctuations, b_i is a time scale parameter related to the initial conditions, k_i is a parameter related to the rate of maturing, and ε_{ij} is the fitting error, which was considered to be independent and normally distributed among individuals. The fitting error variance (σ_ε^2) was considered to be the same for all animals at any time, and thus:

$$f(y_{ij} | a_i, b_i, k_i, \sigma_\varepsilon^2) \sim N\{a_i [1 - b_i \exp(-k_i t_j)]^3, \sigma_\varepsilon^2\}.$$

In a second stage of the hierarchical model, genetic and environmental effects were described in a linear model, explaining the variation of growth curves among individuals. Considering θ as a vector with the parameters \mathbf{a} , \mathbf{b} , and \mathbf{k} for all individuals:

$$\theta = \mathbf{X} \beta + \mathbf{Z} \mathbf{u} + \mathbf{W} \mathbf{c} + \mathbf{e},$$

where β is the vector of environmental effects (contemporary group and age of dam nested in the order of calving); \mathbf{u} is the vector of additive genetic effects; \mathbf{c} is the vector of maternal permanent environmental effects; \mathbf{X} , \mathbf{Z} , and \mathbf{W} are the incidence matrices corresponding to each effect; and \mathbf{e} is the vector of residuals associated with each parameter. These residuals were also considered independent among individuals and normally distributed. However, because of the biological meaning of the parameters, the residual covariances between parameters from the same individual were considered different from zero. Therefore, for each individual i , $\text{cov}(e_{ai}, e_{bi}) \neq 0$, $\text{cov}(e_{ai}, e_{ki}) \neq 0$, and $\text{cov}(e_{bi}, e_{ki}) \neq 0$.

Details about the application of hierarchical models for growth in animal breeding can be found in Varona et al. (1997), Blasco et al. (2003), and Piles et al. (2003).

The Inference

The joint posterior distribution for individual growth curve parameters, their (co)variance components, and environmental and additive genetic effects was estimated under a hierarchical Bayesian framework. Markov Chain Monte Carlo Methods (MCMC) were applied,

using Metropolis-Hastings and Gibbs sampling algorithms (Casella and George, 1992; Chib and Greenberg, 1995; Sorensen and Gianola, 2002).

The Conditional Density of Data. Assuming independence among individuals, the conditional distribution of data \mathbf{y} , given the curve parameters, was a product of normal distributions:

$$f(\mathbf{y} | \theta, \sigma_\varepsilon^2) = \prod_{i=1}^N \prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\sigma_\varepsilon}} \exp \left\{ -\frac{[y_{ij} - a_i(1 - b_i \exp - k_i t_j)^3]^2}{2\sigma_\varepsilon^2} \right\}, \tag{1}$$

where N is the number of individuals with data and n_i is the number of BW recorded on each individual i .

The density of the growth curve parameters, given the genetic and environmental effects, was a multivariate normal distribution:

$$f(\theta | \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}) = |\mathbf{R}|^{-N/2} \exp \left[-\frac{1}{2} (\theta - \mathbf{X}\beta - \mathbf{Z}\mathbf{u} - \mathbf{W}\mathbf{c})' (\mathbf{R} \otimes \mathbf{I})^{-1} (\theta - \mathbf{X}\beta - \mathbf{Z}\mathbf{u} - \mathbf{W}\mathbf{c}) \right], \tag{2}$$

where \mathbf{G} is the additive genetic (co)variance matrix; \mathbf{P} is the maternal permanent environmental (co)variance matrix; \mathbf{R} the residual (co)variance matrix between the parameters a , b , and k ; and \mathbf{I} is an identity matrix.

The Priors. A Bayesian probability model requires assigning prior distributions for all unknown parameters. Normal prior distributions were assumed for the environmental and genetic effects:

$$f(\beta | \mathbf{m}, \mathbf{V}) \propto |\mathbf{V}|^{-1/2} \exp \left[-\frac{1}{2} (\beta - \mathbf{m})' \mathbf{V}^{-1} (\beta - \mathbf{m}) \right],$$

$$f(\mathbf{u} | \mathbf{G}, \mathbf{A}) \propto |\mathbf{G}|^{-N_A/2} \exp \left[-\frac{1}{2} \mathbf{u}' (\mathbf{G} \otimes \mathbf{A})^{-1} \mathbf{u} \right], \text{ and}$$

$$f(\mathbf{c} | \mathbf{P}) \propto |\mathbf{P}|^{-N_d/2} \exp \left[-\frac{1}{2} \mathbf{c}' (\mathbf{I} \otimes \mathbf{P})^{-1} \mathbf{c} \right],$$

where \mathbf{m} and \mathbf{V} are subjective means and (co)variances for the prior beliefs about the systematic effects, N_A is the number of animals in the genealogy, N_d is the number of dams of animals with records, \mathbf{I} is an identity matrix of order N , and \mathbf{A} is the numerator relationship matrix.

Prior distributions for the (co)variance matrices and fitting error variance were assumed to be flat, with limits that guaranteed proper posterior distributions, as follows:

$$\begin{aligned} f(\mathbf{G}) &\propto \text{constant}, \\ f(\mathbf{P}) &\propto \text{constant}, \\ f(\mathbf{R}) &\propto \text{constant, and} \\ f(\sigma_\varepsilon^2) &\propto \text{constant.} \end{aligned}$$

The Posterior Distributions. Bayes theorem, assuming independence between priors, can be expressed by

$$\begin{aligned} f(\theta, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}, \sigma_\varepsilon^2 | \mathbf{y}) &\propto f(\mathbf{y} | \theta, \sigma_\varepsilon^2) \\ &f(\theta | \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}) f(\beta) f(\mathbf{u} | \mathbf{G}) f(\mathbf{G}) \\ &f(\mathbf{c} | \mathbf{P}) f(\mathbf{P}) f(\mathbf{R}) f(\sigma_\varepsilon^2). \end{aligned}$$

The sampling methods require random independent draws from the conditional posterior distribution for each parameter. Supposing θ_{ik} is the k th growth curve parameter for the i th animal, and θ_{-ik} are the other parameters for the i th animal and all parameters for all other animals:

$$\begin{aligned} f(\theta_{ik} | \theta_{-ik}, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}, \sigma_\varepsilon^2, \mathbf{y}) &\propto f(\mathbf{y} | \theta, \sigma_\varepsilon^2) \\ &f(\theta | \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}) \\ &\propto f(\mathbf{y}_i | \theta_i, \sigma_\varepsilon^2) \\ f(\theta_{ik} | \theta_{-ik}, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}) &f(\theta_{-ik} | \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}) \\ &\propto f(\mathbf{y}_i | \theta_{ik}, \theta_{-ik}, \sigma_\varepsilon^2) f(\theta_{-ik} | \theta_{-ik}, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}). \end{aligned}$$

The fully conditional distributions of parameters of hierarchical multistage models were derived by Sorensen and Gianola (2002). In the present work, the fully conditional distributions of growth curve parameters are products of the conditional distribution of data (Eq. [1]) and the prior distribution of the growth curve parameters (Eq. [2]). The same conditional structure was presented by Varona et al. (1999) in a simulation study applying the Von Bertalanffy function.

Supposing \bar{a}_i , \bar{b}_i , and \bar{k}_i are the means of parameters given the genetic and environmental effects, and r^{aa} , r^{bb} , r^{kk} , r^{ab} , r^{ak} , and r^{bk} are the corresponding elements of the inverse of the residual (co)variance matrix (\mathbf{R}), then the fully conditional distribution of parameter a can be written as

$$\begin{aligned} f(a_i | b_i, k_i, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}, \sigma_\varepsilon^2, \mathbf{y}) \\ \propto f(a_i | b_i, k_i, \sigma_\varepsilon^2, \mathbf{y}_i) f(a_i | b_i, k_i, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}), \end{aligned}$$

where

$$f(a_i | b_i, k_i, \sigma_\varepsilon^2, \mathbf{y}_i) \sim N \left[\frac{\sum_{j=1}^{n_i} y_{ij}(1 - b_i \exp(-k_i t_j))^3}{n_i}, \frac{\sigma_\varepsilon^2}{\sum_{j=1}^{n_i} [(1 - b_i \exp(-k_i t_j))^3]^2} \right],$$

and

$$f(a_i | b_i, k_i, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}) \sim N \left[\bar{a}_i - (b_i - \bar{b}_i) \frac{r^{ab}}{r^{aa}} - (k_i - \bar{k}_i) \frac{r^{ak}}{r^{aa}}, \frac{1}{r^{aa}} \right].$$

The fully conditional distribution of parameter b can be written as

$$f(b_i | a_i, k_i, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}, \sigma_\epsilon^2, \mathbf{y}) \propto f(b_i | a_i, k_i, \sigma_\epsilon^2, \mathbf{y}_i) f(b_i | a_i, k_i, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}),$$

where

$$f(b_i | a_i, k_i, \sigma_\epsilon^2, \mathbf{y}_i) \propto \prod_{j=1}^{n_i} \exp \left\{ -\frac{[y_{ij} - a_i(1 - b_i \exp(-k_i t_j))^3]^2}{2\sigma_\epsilon^2} \right\},$$

and

$$f(b_i | a_i, k_i, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}) \sim N \left[\bar{b}_i - (a_i - \bar{a}_i) \frac{r^{ab}}{r^{bb}} - (k_i - \bar{k}_i) \frac{r^{bk}}{r^{bb}}, \frac{1}{r^{bb}} \right].$$

The fully conditional distribution of parameter k can be written as

$$f(k_i | a_i, b_i, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}, \sigma_\epsilon^2, \mathbf{y}) \propto f(k_i | a_i, b_i, \sigma_\epsilon^2, \mathbf{y}_i) f(k_i | a_i, b_i, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}),$$

where

$$f(k_i | b_i, k_i, \sigma_\epsilon^2, \mathbf{y}_i) \propto \prod_{j=1}^{n_i} \exp \left\{ -\frac{[y_{ij} - a_i(1 - b_i \exp(-k_i t_j))^3]^2}{2\sigma_\epsilon^2} \right\},$$

and

$$f(k_i | a_i, k_i, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}) \sim N \left[\bar{k}_i - (a_i - \bar{a}_i) \frac{r^{ak}}{r^{kk}} - (b_i - \bar{b}_i) \frac{r^{bk}}{r^{kk}}, \frac{1}{r^{kk}} \right].$$

Therefore, the parameter a could be easily sampled from a normal distribution, but the conditional posterior densities of parameters b and k did not have a known form. In these cases, the Metropolis-Hastings algorithm with normal proposal distribution centered on the values of b_i and k_i sampled in the immediately previous iteration was used.

The mixed model equations were constructed as if the observed traits were the growth curve parameters:

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} + \mathbf{V}^{-1} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{W} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + (\mathbf{G} \otimes \mathbf{A})^{-1} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{W} \\ \mathbf{W}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{W}'\mathbf{R}^{-1}\mathbf{Z} & \mathbf{W}'\mathbf{R}^{-1}\mathbf{W} + (\mathbf{I} \otimes \mathbf{P})^{-1} \end{bmatrix}$$

$$\begin{bmatrix} \beta \\ \boldsymbol{\mu} \\ \mathbf{c} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\boldsymbol{\theta} \\ \mathbf{Z}'\mathbf{R}^{-1}\boldsymbol{\theta} \\ \mathbf{W}'\mathbf{R}^{-1}\boldsymbol{\theta} \end{bmatrix}.$$

The conditional posterior distributions for each location parameter β_i , \mathbf{u}_i , and \mathbf{c}_i were normal distributions defined by the coefficients and the right-hand side (*RHS*) of the mixed model equations built considering $\boldsymbol{\theta}_{ik}$ as the record:

$$f(\beta_i | \beta_{-i}, \boldsymbol{\theta}, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}, \sigma_\epsilon^2, \mathbf{y}) \sim N \left[\frac{\text{RHS}_i - \sum_{j \neq i} \lambda_{ij} t_j}{\lambda_{ii}}, \frac{1}{\lambda_{ii}} \right],$$

$$f(\mathbf{u}_i | \mathbf{u}_{-i}, \boldsymbol{\theta}, \beta, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}, \sigma_\epsilon^2, \mathbf{y}) \sim N \left[\frac{\text{RHS}_i - \sum_{j \neq i} \lambda_{ij} t_j}{\lambda_{ii}}, \frac{1}{\lambda_{ii}} \right], \text{ and}$$

$$f(\mathbf{c}_i | \mathbf{c}_{-i}, \boldsymbol{\theta}, \beta, \mathbf{u}, \mathbf{G}, \mathbf{P}, \mathbf{R}, \sigma_\epsilon^2, \mathbf{y}) \sim N \left[\frac{\text{RHS}_i - \sum_{j \neq i} \lambda_{ij} t_j}{\lambda_{ii}}, \frac{1}{\lambda_{ii}} \right],$$

where β_{-i} , \mathbf{u}_{-i} , and \mathbf{c}_{-i} are the vectors including the current values of these effects after discarding the i th one, and λ is the corresponding element from the coefficient matrix of the mixed models equations.

Letting n_p be the number of parameters of the growth model, the posterior conditionals of the (co)variance matrices were the following inverted Wishart distributions:

$$\begin{aligned} f(\mathbf{G} | \boldsymbol{\theta}, \beta, \mathbf{u}, \mathbf{c}, \mathbf{P}, \mathbf{R}, \sigma_\epsilon^2, \mathbf{y}) &\sim \text{IW} [(\mathbf{u}'\mathbf{A}^{-1}\mathbf{u}), N_a - (n_p + 1)], \\ f(\mathbf{P} | \boldsymbol{\theta}, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{R}, \sigma_\epsilon^2, \mathbf{y}) &\sim \text{IW} [(\mathbf{c}'\mathbf{c}), N_d - (n_p + 1)], \text{ and} \\ f(\mathbf{R} | \boldsymbol{\theta}, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \sigma_\epsilon^2, \mathbf{y}) &\sim \text{IW} [(\mathbf{e}'\mathbf{e}), N - (n_p + 1)]. \end{aligned}$$

The conditional posterior distribution of the fitting error variance was an inverted χ^2 :

$$f(\sigma_\epsilon^2 | \boldsymbol{\theta}, \beta, \mathbf{u}, \mathbf{G}, \mathbf{c}, \mathbf{P}, \mathbf{R}, \mathbf{y}) \sim \chi^{-2} \left[\sum_{i=1}^N \sum_{j=1}^{n_i} [y_{ij} - a_i(1 - b_i \exp(-k_i t_j))^3]^2, \sum_{i=1}^N n_i - 2 \right].$$

The Chains. Chains of 110,000, 330,000, and 550,000 samples, with different starting values, were carried

out with sampling intervals of 10, 30, and 50, respectively. From each chain, the last 10,000 samples were used to estimate the features of marginal distributions; the burn-in period was greater than the minimum required according to the Raftery and Lewis (1992) method. Convergence was tested for each chain using the criterion of the Geweke (1992) and Gelman and Rubin (1992) test. Autocorrelation between samples and features of marginal distributions of Monte-Carlo error (Geyer, 1992) were calculated.

The Goodness of Fit. The overall goodness of fit was checked by calculating the square of the correlation between observed BW and the predicted value's marginal densities means. A cross-validation predictive density was applied to assess the goodness of fit in different parts of the curve. The observed values, y_{ij} , were compared with their prediction, Y_{ij} , obtained using all of the other data, y_{-ij} . The predictive density of observed data is

$$f(Y_{ij} | y_{-ij}) \propto f(Y_{ij} | \theta, y_{-ij}) f(\theta | y_{-ij}),$$

where y_{-ij} is the vector of observed values discarding the observation being predicted (y_{ij}). It was considered that, if the model was adequate, the observed value, y_{ij} , is a realized random sample from its respective predictive distribution, $f(Y_{ij} | y_{-ij})$, and the evaluation of model fitting was based on the comparison of the observed value with some characteristic of this distribution.

Applying the checking function proposed by Gelfand et al. (1992), $[g(Y_{ij}; y_{ij}) = 1$ if $Y_{ij} < y_{ij}$ and $g(Y_{ij}; y_{ij}) = 0$ if $Y_{ij} \geq y_{ij}]$, the $E[g(Y_{ij}; y_{ij}) | y_{-ij}]$ with respect to $f(Y_{ij} | y_{-ij})$ was computed for each observed record. This expectation showed the probability of a predicted value being lower than the observed one. The average of these expectations for all individuals at each j th point of the growth curve showed the goodness of fit in different parts of the trajectory. If the model fit the data properly, the averages should be close to 0.5. The estimation of $E[g(Y_{ij}; y_{ij}) | y_{-ij}]$ required the computation of the following integral using MCMC methods:

$$E[g(Y_{ij}; y_{ij}) | y_{-ij}] = \iint g(Y_{ij}; y_{ij}) f(Y_{ij} | \theta, y_{-ij}) f(\theta | y_{-ij}) d\theta dY_{ij},$$

which required knowledge of $f(Y_{ij} | \theta, y_{-ij})$ and $f(\theta | y_{-ij})$, which means that resampling all parameters of the growth curve is needed to predict each datum. This would require extremely demanding computing. To avoid this problem, an importance sampling procedure (Gelman et al., 1995) was applied using $f(\theta | y)$ as the sampling distribution for all $f(\theta | y_{-ij})$. If N_{mc} samples of the vector of parameters of the model have been drawn from $f(\theta | y)$, a Monte Carlo estimator of the expectations is

$$\hat{E}[g(Y_{ij}; y_{ij}) | y_{-ij}] = \frac{\sum_{i=1}^{N_{mc}} g(Y_{ij}; y_{ij}) w_{ij}}{\sum_{i=1}^{N_{mc}} w_{ij}}, \text{ and}$$

$$w = \frac{f(y_{-ij} | \theta) f(\theta)}{f(y | \theta) f(\theta)} =$$

$$\frac{\prod_{i=1}^{N-1} \prod_{j=1}^{n_i-1} \frac{1}{\sqrt{2\pi\sigma_\epsilon}} \exp\left\{-\frac{[y_{ij} - a_i(1 - b_i \exp - k_i t_j)^3]^2}{2\sigma_\epsilon^2}\right\}}{\prod_{i=1}^N \prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\sigma_\epsilon}} \exp\left\{-\frac{[y_{ij} - a_i(1 - b_i \exp - k_i t_j)^3]^2}{2\sigma_\epsilon^2}\right\}}, \text{ and}$$

$$w_{ij} = \frac{1}{\frac{1}{\sqrt{2\pi\sigma_\epsilon}} \exp\left\{-\frac{[y_{ij} - a_i(1 - b_i \exp - k_i t_j)^3]^2}{2\sigma_\epsilon^2}\right\}}.$$

The Algorithm. A program was written using Fortran 90. Techniques related to memory use, such as data work file reduction by storing observations, pedigree information, and location parameters on separate files, increased effectiveness of the software by decreasing the reading time from the disk and avoiding storage of large amounts of missing information. The algorithm can be outlined as follows:

1. construct the coefficient matrix;
2. use the fully conditional distribution, sample parameters a , b , and k for all animals;
3. construct the *RHS* of mixed models equations;
4. sample the environmental effects, maternal permanent environmental effects, breeding values, (co)variance components (G , P , and R), and fit the error variance from their respective fully conditional distributions;
5. predict the observed BW of each animal given the growth curve parameters and estimate the Gelfand's checking function for each observation using the importance sampling procedure;
6. repeat steps 2, 3, 4, and 5 for a total of N_{mc} times.

RESULTS AND DISCUSSION

Convergence and Goodness of Fit

Lack of convergence was not detected for any chain. The autocorrelations between samples were low in almost every chain. Heritabilities and genetic correlation estimates showed greater autocorrelations than other parameters, but this did not lead to greater estimates of Monte Carlo SE. Visual inspection of sample trace plots indicated that the sampling processes were appropriate. Features of marginal posterior densities obtained from chains with different lengths, or the same lengths but different starting values, were very similar.

The goodness of fit was considered adequate. The square of the correlation coefficient between the means

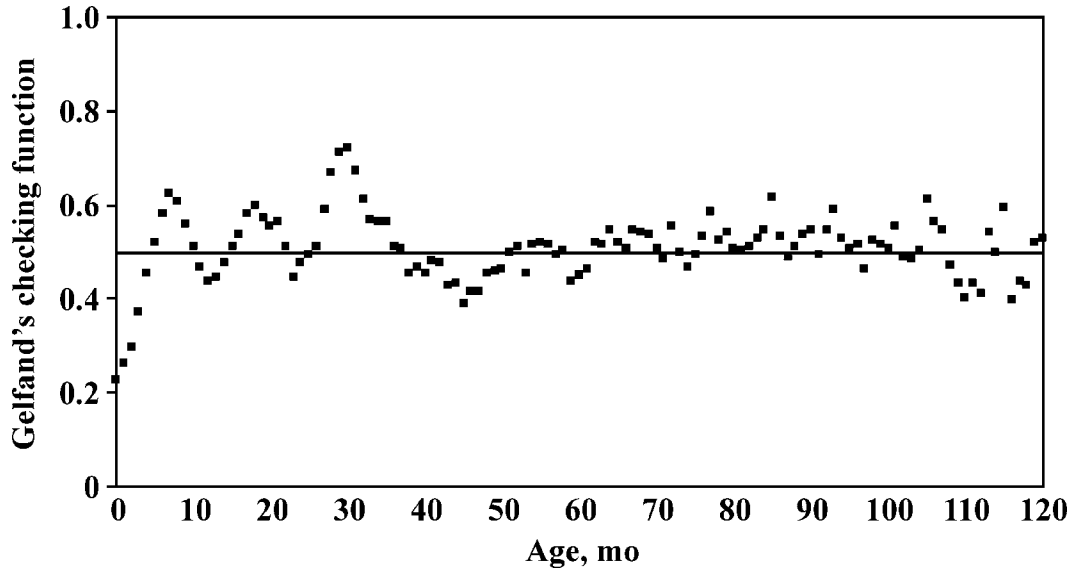


Figure 1. Averages of the expectations of the Gelfand's checking function for all individuals at each point of the Von Bertalanffy growth curve.

of marginal densities of predicted values and observed values was 0.94. Global criteria of fit have the disadvantage of depending more on the last part of the curve than on the first part, due to a scale effect (Blasco et al., 2003). Absolute values of the errors are greater at the adult BW than at the beginning of the growth curve. Because the model could fit better in some parts of the curve than others, and there was a specific interest in the goodness of fit of each parameter, the whole growth trajectory was examined. The probability of a predicted BW being greater or lower than the true value provided a parameter of adjustment quality for each animal at each point of its growth curve. Figure 1 shows the averages of Gelfand's checking function at each measured point of the growth curve. At the beginning, there was a trend of obtaining predictions greater than the observations. The model showed a better fit at the end of the growth curve, after adult BW were reached; the expectations were near 0.5 after maturity was reached; none of the predicted values showed a high probability of being lower than the observed one.

In designing breeding programs for beef cattle, there is particular interest in early predictors of adult BW (Meyer, 1995). When fitting growth curves, the reliability of mature BW is uncertain depending on the data range. Ideally, prediction of mature BW should be based on all available information. The hierarchical procedure applied here provided estimates of the parameters of the growth curve for each animal, taking into account the information obtained from individual performance and the information from relatives and from individuals whose data were affected by the same systematic effects. Even though parameters were also estimated for animals with few records, considering the information of related animals and the structure of systematic effects, mature BW were accurately predicted.

Variance Components

Features of the posterior marginal density of the variance components showed a large additive genetic variance for mature BW (Table 1). An important additive genetic component for this parameter has also been described for beef cattle by DeNise and Brinks (1985), Oliveira et al. (1994), and Meyer (1995). The parameter a of growth curves could be used as a selection criterion to control adult BW increases when selecting for growth rate, mainly in situations when slaughter weight is reached before maturity and adult BW can be recorded for only a few animals. Mature BW was also influenced by maternal permanent environmental effects, probably due to substantial maternal effects on BW until weaning (Albuquerque and Meyer, 2001) and a carryover effect on postweaning growth. Rumph et al. (2002) compared different models for estimation of genetic parameters for mature BW of Hereford cattle and reported that the model that best fit the data included maternal genetic and maternal permanent environmental effects. Archer et al. (1998) compared growth curves from Angus cattle selected for high or low growth rate from birth to a year of age and reported that the effect of maternal environmental on growth was carried through to maturity; low lines calves reared by a high-line dam were heavier and taller at maturity compared with their naturally reared counterparts. Few studies have considered maternal effects when fitting growth curves of beef cattle. Including maternal effects in the models and separating them into genetic and environmental components need to be investigated further.

Parameters b and k showed small genetic and maternal effects. Parameter b is related to the initial conditions of the growth curve, but because birth weight is usually available, it can be used in selection programs

Table 1. Features of the marginal posterior distributions of variance components of Von Bertalanffy growth curve parameters

Parameter	Genetic variance			Maternal permanent environmental variance		
	Mean ± SD	HPD _{95%} ¹	MCse ²	Mean ± SD	HPD _{95%}	MCse
<i>a</i>	864 ± 84	696; 1,028	3.1425	253 ± 44	170; 340	1.2768
<i>b</i> ³	1.56 ± 0.01	1.55 ± 1.58	0.0012	3.12 ± 0.03	3.06; 3.17	0.0033
<i>k</i> ³	0.07 ± 10 ⁻³	0.06; 0.07	0.0009	0.04 ± 10 ⁻³	0.03; 0.04	0.0010
Parameter	Residual variance			Phenotypic variance		
	Mean ± SD	HPD _{95%}	MCse	Mean ± SD	HPD _{95%}	MCse
<i>a</i>	552 ± 42	468; 636	1.3722	1668 ± 62	1,550; 1792	1.6831
<i>b</i> ³	47.3 ± 0.29	46.7; 47.8	0.0118	52.0 ± 0.53	51.0; 53.6	0.0221
<i>k</i> ³	1.12 ± 0.04	1.04; 1.20	0.0016	1.23 ± 0.04	1.16; 1.30	0.0011
Parameter	Heritability			Maternal permanent environmental effects ⁴		
	Mean ± SD	HPD _{95%}	MCse	Mean ± SD	HPD _{95%}	MCse
<i>a</i>	0.52 ± 0.04	0.44; 0.61	0.0037	0.15 ± 0.02	0.11; 0.20	0.0011
<i>b</i>	0.03 ± 0.01	0.01; 0.05	0.0015	0.06 ± 0.02	0.02; 0.09	0.0011
<i>k</i>	0.06 ± 0.01	0.04; 0.09	0.0019	0.03 ± 0.01	0.01; 0.05	0.0007

¹HPD_{95%} = high posterior density interval at a 95% probability.

²MCse = Monte Carlo SE.

³To calculate the actual values, multiply the number by 0.01.

⁴Expressed as the percentage of phenotypic variance.

without estimates of this parameter. Parameter *k* is a function of the ratio of maximum growth rate to mature size (Fitzhugh, 1976) and could be used as a selection criterion indicating the rate of approach to mature BW. However, due to its low additive genetic variance, its inclusion in selection programs would not be effective. Large environmental effects on parameter *k* were also reported in studies applying the Von Bertalanffy function in a 2-step analysis of growth curves (Oliveira et al., 1994).

Covariance Components

Covariance estimates showed a negative genetic correlation between mature BW and parameter *b* (Table 2). This result was expected because greater values of parameter *b* lead to lower birth weights [birth weight = $a \cdot (1 - b)^3$] and it is known that BW at different ages

show a large positive genetic correlation (Albuquerque and Meyer, 2001). Greater values of adult BW are associated with greater BW in all periods of growth. DeNise and Brinks (1985) also reported a negative genetic correlation between the corresponding *a* and *b* parameters of the Brody growth curve in beef cattle. These parameters are similarly interpreted in Brody and Von Bertalanffy curves.

Parameters *a* and *k* were estimated to have a large positive genetic correlation, with changes in both parameters thus being expected as a consequence of selection for growth rate. However, the additive genetic variance of *k* was very low; thus, a large correlated response to selection is not expected. Archer et al. (1998) observed no responses in the rate of maturation of cows and steers selected for different growth rates and concluded that there was no change in the temporal pattern of maturation despite differences in mature

Table 2. Features of the marginal posterior distributions of correlations between Von Bertalanffy growth curve parameters

Parameters	Genetic correlation			Maternal permanent environmental correlation		
	Mean ± SD	HPD _{95%} ¹	MCse ²	Mean ± SD	HPD _{95%}	MCse
<i>a</i> _b	-0.38 ± 0.14	-0.68; -0.11	0.0182	-0.18 ± 0.12	-0.42; 0.06	0.0071
<i>a</i> _k	0.82 ± 0.07	0.71; 0.95	0.0087	0.46 ± 0.18	0.12; 0.82	0.0151
<i>b</i> _k	-0.24 ± 0.18	-0.58; 0.09	0.0229	0.07 ± 0.18	-0.31; 0.40	0.0116
Parameters	Phenotypic correlation			Residual correlation		
	Mean ± SD	HPD _{95%}	MCse	Mean ± SD	HPD _{95%}	MCse
<i>a</i> _b	-0.02 ± 0.04	-0.09; 0.05	0.0030	0.07 ± 0.05	-0.02; 0.17	0.0028
<i>a</i> _k	0.28 ± 0.04	0.20; 0.36	0.0039	0.18 ± 0.05	0.08; 0.28	0.0033
<i>b</i> _k	0.66 ± 0.02	0.62; 0.69	0.0014	0.73 ± 0.02	0.68; 0.78	0.0019

¹HPD_{95%} = high posterior density interval at a 95% probability.

²MCse = Monte Carlo SE.

size. Therefore, selection for growth rate would increase adult BW without substantially changing the shape of the growth curve as predicted by Taylor (1985). There has been much interest over the years in the potential for bending the growth curve to combine high growth rates with small mature size to improve efficiency of animal production (Fitzhugh, 1976). However, genetic changes in the shape of the growth curve are limited by the dependence between mature size and rate of maturing. The large positive genetic correlation between the parameters a and k , combined with the positive genetic correlation between all BW along the growth curve (Lôbo et al., 2000; Albuquerque and Meyer, 2001), suggest that a selection program to change the slope of the growth curve without increasing adult BW would be inefficient. Decreasing growth rates could be observed as a correlated response to selection, if boundaries on adult BW were established.

The length of the HPD_{95%} of the genetic correlation between parameters b and k was too large to permit reliable conclusions, but these parameters clearly showed large positive phenotypic and residual correlations. Regardless of that, the relationship between the initial conditions and subsequent growth is explained well by the large and positive genetic correlations between early BW.

The correlation between maternal permanent environmental effects of parameters a and k suggested that, besides additive genetic effects, variations observed in both parameters in the same direction could also be caused by these effects. Maternal environment may have an overall influence on growth curves, but its relevance for parameters b and k was not clear. Optimum mature size for beef cattle is commonly debated. Cattle vary widely in body size, and the optimum depends on the production system. In general, selection programs have placed emphasis on greater growth rates. This led to changes in body size at all ages, which may be evaluated under specific production characteristics.

In the present work, adjustment error variance was considered to be constant along the curve, when obviously it should be smaller at the beginning of the growth process due to a scale effect. Growth data commonly present time-series errors, which were partially overlooked in this study. Considering other error structures leads to both overparameterization of the model and lack of convergence of the MCMC chains, as previous analyses of the data showed. Using less complex error structures is a usual procedure to avoid these problems, although it leads to some underestimation of the variance components of the parameters (Glasbey, 1979).

Mature BW, predicted by parameter a of the growth curves, could be included in selection programs. This parameter could be predicted by combining individual BW with information derived from the parents' mature BW, as well as BW of other relatives. Selection to change the slope of the growth curve without modi-

fying adult BW would be inefficient because their genetic correlation is large. However, adult BW could be considered in a selection index with its corresponding economic weight to improve the overall efficiency of beef cattle production.

LITERATURE CITED

- Albuquerque, L. G., and K. Meyer. 2001. Estimates of covariance functions for growth from birth to 630 days of age in Nelore cattle. *J. Anim. Sci.* 79:2776–2789.
- Arango, J. A., and L. D. Van Vleck. 2002. Size of beef cows: Early ideas, new developments. *Genet. Mol. Res.* 1:51–63.
- Archer, J. A., R. M. Herd, P. F. Arthur, and P. F. Parnell. 1998. Correlated responses in rate of maturation and mature size of cows and steers to divergent selection for yearling growth rate in Angus cattle. *Livest. Prod. Sci.* 54:183–192.
- Blasco, A., M. Piles, and L. Varona. 2003. A Bayesian analysis of the effect of selection for growth rate on growth curves in rabbits. *Genet. Sel. Evol.* 35:21–41.
- Casella, G., and E. I. George. 1992. Explaining the Gibbs Sampler. *Am. Stat.* 46:167–174.
- Chib, S., and E. Greenberg. 1995. Understanding the Metropolis-Hastings algorithm. *Am. Stat.* 49:327–335.
- DeNise, R. S. K., and J. S. Brinks. 1985. Genetic and environmental aspects of the growth curve parameters in beef cows. *J. Anim. Sci.* 61:1431–1440.
- Fitzhugh, H. A., Jr. 1976. Analysis of growth curves and strategies for altering their shape. *J. Anim. Sci.* 42:1036–1051.
- Gelfand, A., D. K. Dey, and H. Chang. 1992. Model determination using predictive distributions with implementation via sampling-based methods (with discussion). Pages 147–167 in *Bayesian Statistics 4*. Oxford Univ. Press, Oxford, UK.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 1995. *Bayesian Data Analysis*. Chapman & Hall, London, UK.
- Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7:457–472.
- Geweke, J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments (with discussion). Pages 169–193 in *Bayesian Statistics 4*. Oxford Univ. Press, Oxford, UK.
- Geyer, C. J. 1992. Practical Markov Chain Monte Carlo. *Stat. Sci.* 7:473–511.
- Glasbey, C. A. 1979. Correlated residuals in nonlinear regression applied to growth data. *Appl. Stat.* 28:251–259.
- Lôbo, R. B., L. A. F. Bezerra, H. N. Oliveira, C. U. Magnabosco, A. R. Zambianchi, L. G. Albuquerque, J. A. G. Bergmann, and R. D. Sainz. 2005. *Avaliação Genética de Touros e Matrizes da Raça Nelore*. 11th ed. Associação Nacional de Criadores e Pesquisadores, Ribeirão Preto, SP, Brazil.
- Lôbo, R. N. B., F. E. Madalena, and A. R. Vieira. 2000. Average estimates of genetic parameters for beef and dairy cattle in tropical regions. *Anim. Breed. Abstr.* 68:433–462.
- Meyer, K. 1995. Estimates of genetic parameters for mature weight of Australian beef cows and its relationship to early growth and skeletal measures. *Livest. Prod. Sci.* 44:125–137.
- Mignon-Gastreau, S., M. Piles, L. Varona, H. De Rochambeau, A. Blasco, and C. Beaumont. 2000. Bayesian analysis of Gompertz curve of chickens selected on shape of growth curve. *J. Anim. Sci.* 78:2515–2524.
- Oliveira, H. N., R. B. Lôbo, and C. S. Pereira. 1994. Relationship among growth curve parameters, weights and reproductive traits in Guzerá beef cows. Page 19 in *Proc. 5th World Cong. Genet. Appl. Livest. Prod.*, Guelph, Canada.
- Piles, M., D. Gianola, L. Varona, and A. Blasco. 2003. Bayesian inference about parameters of a longitudinal trajectory when selection operates on a correlated trait. *J. Anim. Sci.* 81:2714–2724.

- Raftery, A. E., and S. Lewis. 1992. How many iterations in the Gibbs sampler? Pages 763–773 in *Bayesian Statistics 4*. Oxford Univ. Press, Oxford, UK.
- Rumph, J. M., R. M. Koch, K. E. Gregory, L. V. Cundiff, and L. D. Van Vleck. 2002. Comparison of models for estimation of genetic parameters for mature weight of Hereford cattle. *J. Anim. Sci.* 80:583–590.
- Sorensen, D., and D. Gianola. 2002. *Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics*. Springer-Verlag, London, UK.
- Taylor, S. T. C. S. 1985. Use of genetic size-scaling in evaluation of animal growth. *J. Anim. Sci.* 61:118–143.
- Varona, L., C. Moreno, L. A. García Cortés, and J. Altarriba. 1997. Multiple trait genetic analysis of underlying biological variables of production functions. *Livest. Prod. Sci.* 47:201–209.
- Varona, L., C. Moreno, L. A. García Cortés, G. Yangue, and J. Altarriba. 1999. Two-steps versus joint analysis of Von Bertalanffy function. *J. Anim. Breed. Genet.* 116:331–338.
- Von Bertalanffy, L. 1957. Quantitative laws in metabolism and growth. *Q. Rev. Biol.* 32:217–230.

References

This article cites 20 articles, 6 of which you can access for free at:
<http://jas.fass.org/cgi/content/full/85/12/3189#BIBL>

Citations

This article has been cited by 1 HighWire-hosted articles:
<http://jas.fass.org/cgi/content/full/85/12/3189#otherarticles>