



ELSEVIER

Livestock Production Science 93 (2005) 213–222

**LIVESTOCK
PRODUCTION
SCIENCE**

www.elsevier.com/locate/livprodsci

Estimates of covariance functions for growth of Nelore cattle applying a parametric correlation structure to model within-animal correlations

L.G. Albuquerque^{*,1}, K. Meyer

Animal Genetics and Breeding Unit, University of New England², Armidale NSW 2351, Australia

Received 3 April 2003; received in revised form 27 April 2004; accepted 11 October 2004

Abstract

A total of 20,065 weights recorded on 3016 Nelore animals were used to estimate covariance functions for growth from birth to 630 days of age, assuming a parametric correlation structure to model within-animal correlations. The model of analysis included fixed effects of contemporary groups and age of dam as quadratic covariable. Mean trends were taken into account by a cubic regression on orthogonal polynomials of animal age. Genetic effects of the animal and its dam and maternal permanent environmental effects were modelled by random regressions on Legendre polynomials of age at recording. Changes in direct permanent environmental effect variances were modelled by a polynomial variance function, together with a parametric correlation function to account for correlations between ages. Stationary and nonstationary models were used to model within-animal correlations between different ages. Residual variances were considered homogeneous or heterogeneous, with changes modelled by a step or polynomial function of age at recording. Based on Bayesian information criterion, a model with a cubic variance function combined with a nonstationary correlation function for permanent environmental effects, with 49 parameters to be estimated, fitted best. Modelling within-animal correlations through a parametric correlation structure can describe the variation pattern adequately. Moreover, the number of parameters to be estimated can be decreased substantially compared to a model fitting random regression on Legendre polynomial of age.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Covariance functions; Nelore cattle; Parametric correlation structure

1. Introduction

Regressions on Legendre polynomials of age have been used extensively to model longitudinal data for traits such as milk yields and growth. In general, such models involve high-order polynomials, especially for animal permanent environmental effects (Brother-

* Corresponding author. Tel.: +55 16 3209 2678; fax: +55 16 3202 4275.

E-mail address: lgalb@fcav.unesp.br (L.G. Albuquerque).

¹ On leave from: Departamento de Zootecnia, FCAV, UNESP, 14884-900 Jaboticabal, SP, Brazil.

² AGBU is a joint venture of the University of New England and NSW Agriculture.

stone et al., 2000; Meyer, 2001a; Albuquerque and Meyer, 2001) and, consequently, a large number of parameters to be estimated. Higher order polynomials can be associated with sampling problems (Kirkpatrick et al., 1994; Meyer, 1998a) and large computational requirements.

Alternatively, within-animal (co)variation could be modelled through a variance function combined with a parametric correlation function (Pletcher and Geyer, 1999; Foulley et al., 2000). Most commonly, parametric correlation functions assume stationarity, i.e., that correlations depend on the difference in time between records, usually referred to as lag, and have only one or two parameters to be estimated (Nuñez-Anton and Zimmerman, 2000). Pletcher and Geyer (1999) called this model “character process model,” and an extension to nonstationary correlations and comparisons with random regression models are given by Jaffrézic and Pletcher (2000). Software packages such as SAS (1997) and ASREML (Gilmour et al., 1999) can be used to fit parametric correlation functions; see Littell et al. (2000) and Jaffrézic et al. (2002) for applications.

The use of parametric correlation functions in animal breeding context so far has been limited. First-order autoregressive [AR(1)] models have been proposed to model random effects in analyses of dairy cattle data such as permanent environmental effects (Harville, 1979; Mansour et al., 1985; Carvalheira et al., 2002) and contemporary groups (Wade et al., 1993). Models with stationary correlations but allowing for changes in variances with age have also been developed (Mansour et al., 1985; Pletcher and Geyer, 1999; Jaffrézic et al., 2002). Meyer (2001b) presented a set of models combining random regression on Legendre polynomials of age to model variances between animals and a variance function associated to a parametric correlation function, stationary and nonstationary, to model within-animal (co)variances. Applying those models to a data set of mature weight of beef cows, Meyer (2001b) showed that assuming a parametric correlation structure for within-animal covariances can result in a more parsimonious model than a random regression model for permanent environmental effects. Bermejo et al. (2003), working with daily feed intake in pigs, used a quadratic random regression model for additive genetic effects and different combinations of covari-

ariance structures to model environmental effects. Based on Akaike’s Information Criterion, the authors concluded that the parsimonious model included a constant diagonal covariance structure for permanent environmental effects and a heterogeneous autoregressive-moving-average structure for temporary environmental effects.

The objective of the present work was to estimate genetic covariance functions for growth of Nelore cattle using random regression models, assuming a parametric correlation structure to describe animal permanent environmental effects as proposed by Foulley et al. (2000).

2. Material and methods

2.1. Data

Nelore is a Zebu breed adapted to extensive production systems in tropical environment and represents 80% of Brazilian beef cattle population. The data used for this study has been analysed previously, and further details are given by Albuquerque and Meyer (2001). A total of 20,065 weights recorded on 3016 Nelore animals, which were offspring of 87 sires and 1903 dams, and distributed in 523 contemporary groups (CG) in a single herd were used. CG were defined as year and month of birth, sex, weaning status (suckling or weaned) and year and month of weighing subclasses.

Animals were weighed every 3 months from birth to 630 days of age. Only data from animals raised on pasture, in CG with a minimum of four animals and with at least three records were kept. All animals had birth weight, and 79% of the animals had six records or more.

Pedigree information was obtained for up to three generations back, resulting in a total of 5751 animals in the relationship matrix. A summary of the data structure is presented in Table 1. In order to reduce the number of age combinations and decrease computational demand, animals weighed within 4 days were considered the same age, creating 156 age classes, including birth weight as a separate class. No animals were weighed between the 1st and the 12th day of age. Mean weights and number of animals for 4 days interval are presented in Fig. 1.

Table 1
Summary of the data structure

	Number
Records	20,065
Animals with records	3016
Sires	87
Dams	1903
Animals in relationship matrix	5751
Contemporary groups	523
Parameters	33–62

2.2. Analysis

Correlation and (co)variance functions were estimated by restricted maximum likelihood (REML) using the program DXMRR (Meyer, 1998b). Fixed effects fitted were CG and age of dam as a quadratic

covariable. Mean trends were taken into account by a cubic regression on orthogonal polynomials of animal age. Genetic effects of animal and dam and maternal permanent environmental effects were modelled by random regressions on Legendre polynomials of age at recording. Changes in animal permanent environmental effect variances were modelled by a polynomial variance function (VF) combined with a parametric correlation function (RF) to account for correlations between ages. The model can be represented in matrix notation as

$$y = \mathbf{XB} + \mathbf{Z}_1\alpha + \mathbf{Z}_2\gamma + \mathbf{W}_1\delta + \mathbf{W}_2\rho + e$$

where y is the vector of observations, B the vector of fixed effects, α and γ , the vectors of additive genetic direct and maternal random coefficients, δ and ρ , the

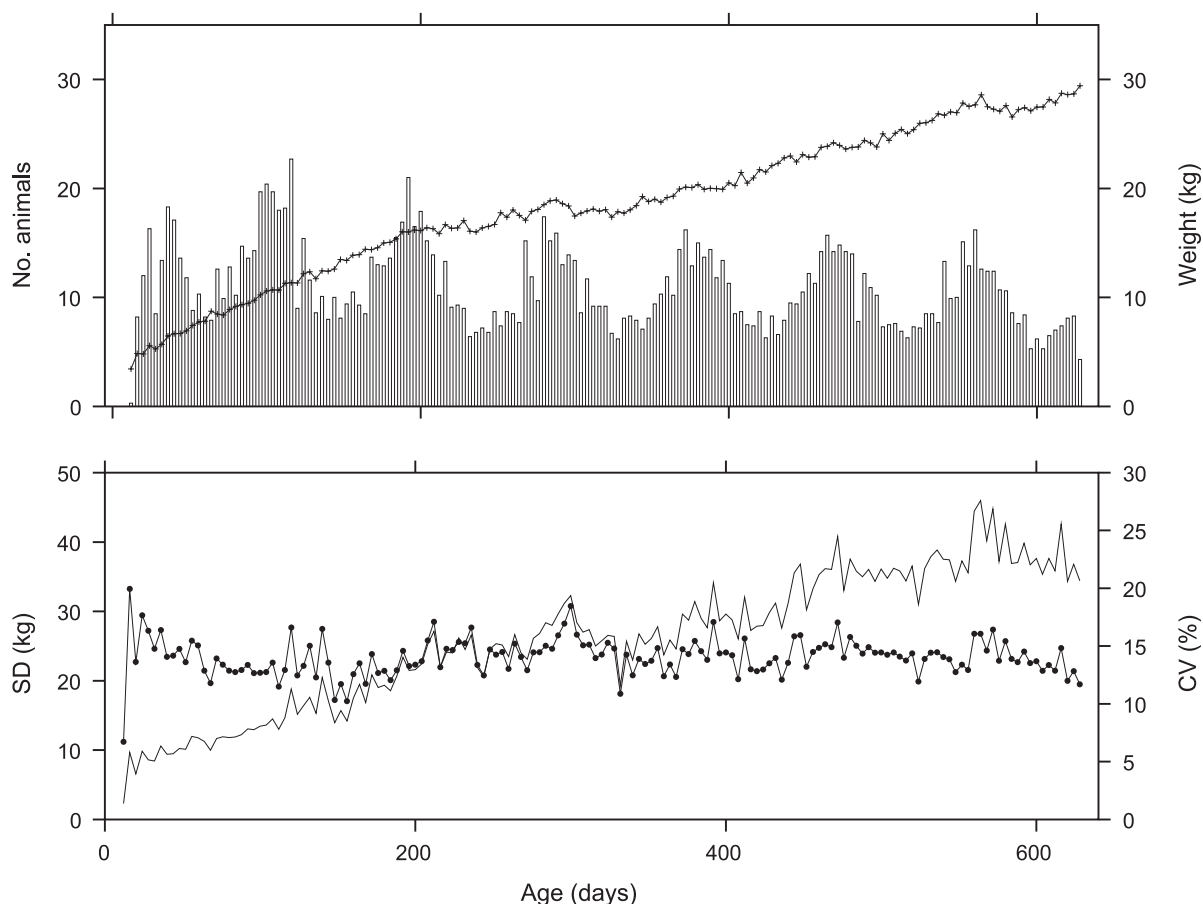


Fig. 1. Top: Weight means divided by 10 (+) and number of observations divided by 10 (bars). Number of animals at birth not included. Bottom: standard deviations (●) and coefficient of variation (solid line).

vectors of direct and maternal permanent environmental random coefficients, \mathbf{X} , \mathbf{Z}_1 , \mathbf{Z}_2 , \mathbf{W}_1 , and \mathbf{W}_2 are the correspondent incidence matrices, and \mathbf{e} is the vector of temporary environmental effects. It was assumed that $\mathbf{V}(\alpha)=\mathbf{K}_a \otimes \mathbf{A}$, $\mathbf{V}(\gamma)=\mathbf{K}_m \otimes \mathbf{A}$, $\mathbf{V}(\delta)=\mathbf{Q}$, $\mathbf{V}(\rho)=\mathbf{K}_c \otimes \mathbf{I}$, and $\mathbf{V}(\mathbf{e})=\mathbf{R}$, and that covariances between different effects were zero. \mathbf{K}_a , \mathbf{K}_m , and \mathbf{K}_c are the matrices of (co)variances between the regression coefficients for additive direct genetic and maternal genetic effects and maternal permanent environmental effects, respectively, \mathbf{A} is the additive numerator relationship matrix; \mathbf{I} is an identity matrix, and \mathbf{Q} is the matrix of covariances for animal permanent environmental effects. The latter can be decomposed as

$$\mathbf{Q} = \Sigma_q^{1/2} \mathbf{C} \Sigma_q^{1/2}$$

with $\Sigma_q = \text{Diag}\{\sigma_{qj}^2\}$ the diagonal matrix of permanent environmental variances and $\mathbf{C} = \{c_{jk}\}$ the correspondent matrix of correlations. \mathbf{C} is block diagonal for animals; see Meyer (2001b) for details. As in previous analyses (Albuquerque and Meyer, 2001), polynomials order of fit for maternal genetic (k_m), and permanent environmental (k_c) effects were 4 and 3, respectively. Polynomial orders of 6 or 4 were fitted for genetic direct (k_a) effects.

Temporary environmental effects (\mathbf{e}) were considered independently distributed, assuming homogeneity or heterogeneity of variances. Heterogeneity of variances was modelled through a step function, with six classes (0–103, 104–159, 160–255, 256–527, 528–591, and 592–630 days) or seven classes (adding a separate class for birth weight), based on residual variance estimates from previous analyses (Albuquerque and Meyer, 2001). In addition, residual variances in the log scale were modelled through a cubic variance function (VF), as in Albuquerque and Meyer (2001).

2.2.1. Permanent environmental variance functions

Changes in permanent environmental variances due to animal with age at recording were modelled by a polynomial variance function as $\sigma_j^2 = \sigma_0^2 \{1 + \sum_{r=1}^v b_r (a_j^*)^r\}$, where σ_j^2 is the variance at the j th age, σ_0^2 is the variance at the mean age, b_r and v are the coefficients and the

number of parameters of the VF, respectively, and a_j^* is the standardised (–1 to +1) age at recording. Preliminary analyses showed no advantage in modelling direct permanent environmental variances by a log function. Orders of fit for VF varied from 2 to 6.

2.2.2. Correlation functions

For stationary correlation functions, the correlations between weights at any two ages are a function of the difference in age at recording. Both stationary and nonstationary correlation functions were used to model within-animal correlations between repeated records. Formulae for the correlation functions are presented in Table 2. Three stationary correlation functions (EXP: exponential; GAU: Gaussian; and DEX: “damped exponential”) and a nonstationary function (SAD: first-order structured antedependence model) were considered, see Diggle et al. (1994) and Nuñez-Anton and Zimmerman (2000).

EXP has a single parameter to be estimated (θ), the exponential equivalent to an autocorrelation. When the decrease in correlations with the increase in distance between ages is faster than that modelled by EXP, GAU, which differs from EXP by using the square of the lag in ages, may be more appropriated. DEX has a second parameter (κ) to be estimated and is more flexible than both the above models. κ is a scale parameter which permits attenuation or acceleration of the exponential decay of the autocorrelation function. These three functions can be represented as

Table 2
Correlation functions

Correlation Function	Parameters	r_{ij}
Exponential (EXP)	θ	$e^{-\theta(t_j - t_i)}$
Gaussian (GAU)	θ	$e^{-\theta(t_j - t_i)^2}$
“Damped” exponential (DE)	θ, κ	$e^{-\theta(t_j - t_i)^\kappa}$
Structured antedependence (SAD)	ρ, κ	$\rho^{f(t_j, \kappa) - f(t_i, \kappa)}$ for $j = i + 1$
		$\prod_{k=i+1}^{j-1} r_{k(k+1)}$ for $j > i + 1$

r_{ij} : correlation between records taken at time t_i and t_j , $i < j$; ρ : autocorrelation parameter; θ : corresponding exponential parameter; and κ : scale parameter.

$(\rho^{(t_j-t_i)})$ where $\rho=e^{-\theta}$ with $k=1,2$, κ , respectively for EXP, GAU, and DEX.

The structured antedependence model, SAD, proposed by Nuñez-Anton and Zimmerman (2000), is a nonstationary correlation function. To accommodate nonstationarity, SAD applies a Box–Cox transformation to the time scale such as $f(t_i, \kappa)=(t_i^k-1)/\kappa$ for $\kappa \neq 0$ and $\log(t_i)$ for $\kappa=0$. Equidistant correlations will vary with age if $\kappa \neq 1$, increasing if $\kappa < 1$, and decreasing if $\kappa > 1$ (Nuñez-Anton and Zimmerman, 2000).

Models were compared using the log likelihood function ($\log L$), REML forms of Akaike's (AIC) and Schwarz' Bayesian (BIC) information criterion (see Wolfinger, 1993) and by examining variance and correlation estimates for ages in the data. With AIC and BIC, the log likelihood is penalised for the number of parameters to be estimated, and non-nested models can be compared. The information criteria formulas are

$$AIC = -2\log L + 2p, \text{ and}$$

$$BIC = -2\log L + p\log(N - r_X)$$

where p is the number of parameters estimated, N is the sample size, r_X is the rank of the coefficient matrix for fixed effects.

Results were contrasted to estimates from previous analyses (Albuquerque and Meyer, 2001), fitting Legendre polynomials of age to model all four random effects, i.e., direct and maternal genetic and permanent environmental effects. In the following, "LP k " denotes a model fitting a Legendre polynomial of order k for animal permanent environmental effects, and "X.VF v " is a model fitting a parametric correlation function together with a variance function of order v , with X=EXP, GAU, DEX or SAD.

3. Results

Results are summarized in Table 3. Estimates of correlations between records one age class apart, ρ , were high for all models. For models fitting $k_a=4$ and a VF for residual variances, estimates of κ were close to unity, reducing SAD to an autocorrelation function. In this case, there were small differences in $\log L$, and BIC was smaller for EXP than for DEX and SAD.

Table 3

Number of parameters (p), log-likelihood function ($\log L+51,500$), Akaike's (AIC) and Schwarz' Bayesian (BIC) information criterion (both $-103,000$) and estimates of correlation function parameters^a for analyses fitting different models

Model	p	$\log L$	AIC	BIC	ρ	θ	κ
<i>k_a=6 and residual modelled by cubic variance function</i>							
LP6	62	-31	187	677	1.000*		
EXP.VF3	46	-43	178	541	0.999	0.011	1.000**
GAU.VF3	46	-88	268	630	0.998	0.002	2.000**
DEX.VF2	46	-62	216	580	0.899	0.011	0.907
DEX.VF3	47	-42	177	547	0.978	0.022	0.778
DEX.VF4	48	-41	179	558	0.977	0.023	0.766
SAD.VF3	47	-30	154	524	0.866	0.144	0.290
SAD.VF4	48	-30	156	535	0.862	0.148	0.283
SAD.VF5	49	-30	158	544	0.865	0.145	0.281
SAD.VF6	50	-28	156	550	0.866	0.144	0.281
<i>k_a=6 and step function for residuals</i>							
EXP.VF3.e6	48	-14	125	505	0.988	0.012	1.000**
SAD.VF3.e1	44	-61	224	626	0.944	0.057	0.840
SAD.VF3.e6	49	-5	108	494	0.956	0.045	0.616
SAD.VF3.e7	50	-2	103	497	0.908	0.097	0.419
<i>k_a=4 and residual modelled by cubic variance function</i>							
LP6	51	-61	224	625	1.000*		
EXP.VF3	35	-125	320	597	0.981	0.019	1.000**
DEX.VF3	36	-124	319	603	0.983	0.017	1.022
DEX.VF4	37	-123	319	611	0.985	0.015	1.015
DEX.VF5	38	-121	318	617	0.983	0.017	1.002
SAD.VF3	36	-122	317	600	0.980	0.020	0.949
<i>k_a=4 and step function for residuals</i>							
EXP.VF3.e7	38	-75	225	526	0.986	0.014	1.000**
SAD.VF3.e1	33	-145	356	616	0.982	0.019	1.044
SAD.VF3.e6	38	-75	226	526	0.945	0.057	0.587
SAD.VF3.e7	39	-63	206	513	0.926	0.765	0.465

^a ρ : correlation parameter for SAD (corresponding value calculated for EXP and DEX as $e^{-\theta}$); θ : exponential parameter for EXP, GAU and DEX (corresponding value calculated for SAD as $-\log(\rho)$); κ : scaling parameter for DEX and SAD.

* Average of correlations between records taken one day apart, see text for explanation.

** Values of κ for EXP and GAU as previously defined.

However, for models fitting $k_a=4$ and 6 or 7 classes for residual variances, estimates of κ were close to 0.5 and for models fitting $k_a=6$, κ varied from 0.28 to 0.84, i.e., correlations between records equidistant in time increased with age (Nuñez-Anton and Zimmerman, 2000).

Increasing the order of fit for the VF for permanent environmental effects from 2 to 3 increased $\log L$ significantly and decreased AIC and BIC. However, regardless of the order of fit for the other

effects in the model or the parametric correlation function utilised (DEX or SAD), VF orders of fit higher than 3 did not change $\log L$ significantly. Considering the same order of fit for all other effects in the model, every RF model resulted in smaller BIC values than the LP model, even for homogeneous residual variance. This was due to the stringent penalty for the number of parameters inherent to BIC. Modelling PE variances through a RF instead of a LP reduced the number of parameters by at least 12.

Based on both criteria, AIC and BIC, a nonstationary model was considered the best. Meyer (2001b), working with a very small data set on mature weights of beef cows, did not find any advantage of SAD over DEX. Based on BIC, a model fitting a quintic polynomial ($k_a=6$) for animal genetic effects, SAD.VF3 and six classes for residual variances, with 49 parameters to be estimated, was the best. The smallest AIC value was for a model with the same specification as above but with seven classes for residual variances. However, the estimate of residual variance for the last class was extremely high, affecting all the parameter estimates. Based on AIC and BIC, when animal genetic effects were modelled by a cubic polynomial

regression ($k_a=4$), a model with SAD.VF3 and seven classes for residual variances (39 parameters to be estimated) fitted best.

3.1. Variance component estimates

Variance component estimates obtained using the two Legendre models and models fitting SAD are presented in Fig. 2. In general, variance estimates were close to those estimated considering the corresponding LP model. For $k_a=6$, changing the residual from six classes (49 parameters) to a cubic VF (47 parameters) did not affect the variance estimates notably. Moreover, estimates obtained with the latter model were even slightly closer to those obtained using the corresponding model fitting LP, although differences in $\log L$ were relatively large. These results agree with those from Olori et al. (1999), who found that different models for residual variances can produce large changes in $\log L$ and in estimates of residual variances without affecting estimates of genetic direct variance components. However, when direct genetic effects were modelled by a cubic regression on Legendre polynomials of age ($k_a=4$), changing the residual from seven to six classes, or to a cubic variance function, produced large changes in the

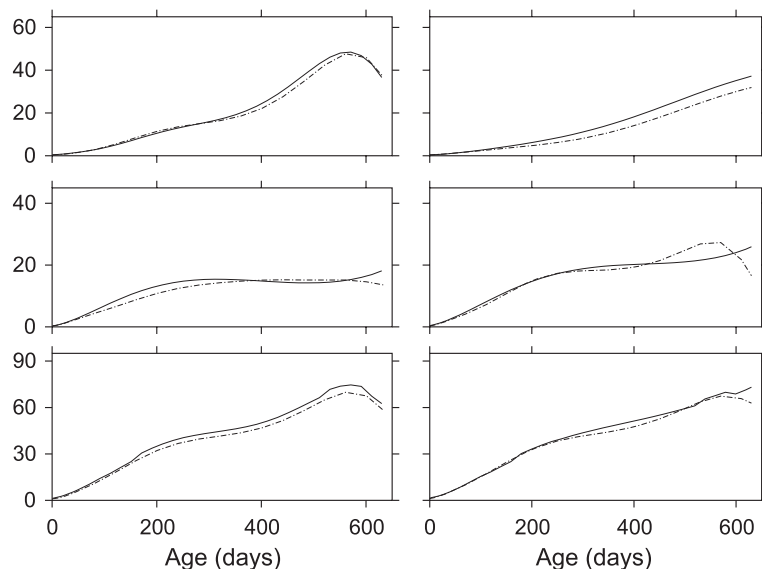


Fig. 2. Top: genetic direct variances. Middle: direct permanent environmental variances. Bottom: phenotypic variances. All variances divided by 10. Left: $k_a=6$, SAD.VF3 and six classes for residual (solid line) and $k_a=6$, LP6 and a cubic variance function for residual (dash-dot line). Right: $k_a=4$, SAD.VF3 and seven classes for residual (solid line) and $k_a=4$, LP6 and a cubic variance function for residual (dash-dot line).

direct variance component estimates. The variance component estimates were more sensitive to changes in the residual when the genetic direct effect was modelled by a lower order polynomial (cubic instead of a quintic).

Differences between models fitting SAD parametric correlation function and the corresponding Legendre model occurred mainly in the partitioning of direct genetic and permanent environmental variances. Maternal variance component estimates obtained with SAD were virtually the same as those obtained with Legendre polynomials and were not affected by changes in the residual model (not shown).

3.2. Heritability and correlation estimates

In general, for the best models, direct heritability estimates (Fig. 3) obtained using a parametric correlation function to model permanent environmental effects were close to those estimated using Legendre polynomials. For $k_a=4$, estimates obtained using SAD were closer to those obtained using standard univariate analysis for weights at every month of age (Albuquerque and Meyer, 2001) than the corresponding LP model. Differences in direct heritability estimates were accompanied by differ-

ences in residual and animal permanent variances as proportions of phenotypic variances but did not affect maternal genetic and permanent environmental estimates (not shown).

To illustrate the effect of different parametric correlation models on phenotypic and direct genetic and permanent environmental correlation estimates, Fig. 4 presents average estimates for the lag in ages between records. EXP, GAU and DEX are stationary functions, i.e., correlations are assumed to depend only on lag in age between records. Direct permanent environmental correlation estimates obtained by EXP and GAU showed a steady decline with increase in lag. This decline was not so steep for DEX since this function has a second parameter (κ) which attenuates the exponential decay when κ is different from unity. LP and SAD are nonstationary models, and average of correlations between records taken one day apart for animal permanent environmental effects are presented. The decrease in within-animal correlation estimates with lag in age at recording was slower than for the stationary functions.

The model assumed for animal permanent environmental effects practically did not affect phenotypic correlation estimates. As observed by Albuquerque and Meyer (2001), maternal correlations were con-

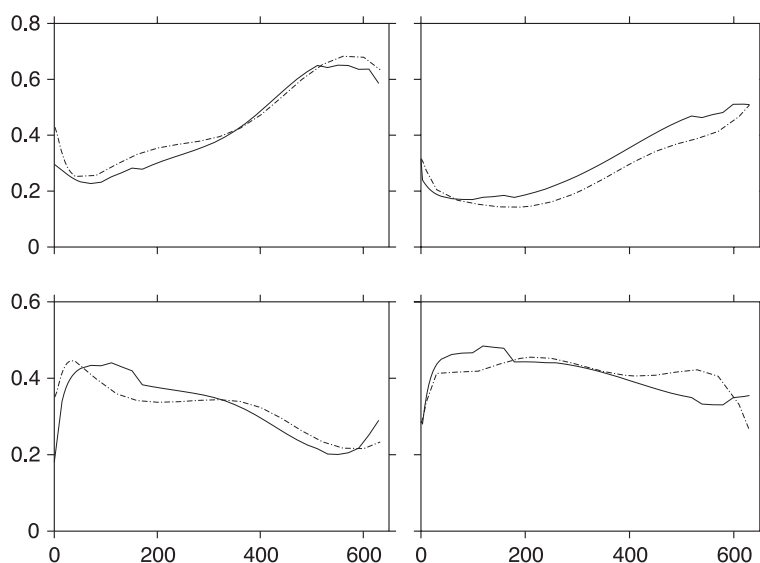


Fig. 3. Top: direct heritability estimates. Bottom: estimates of animal permanent environmental effect variances as proportion of phenotypic variances. Left: $k_a=6$, SAD.VF3 and six classes for residual (solid line) and $k_a=6$, LP6 and a cubic variance function for residual (dash-dot line). Right: $k_a=4$, SAD.VF3 and seven classes for residual (solid line) and $k_a=4$, LP6 and a cubic variance function for residual (dash-dot line).

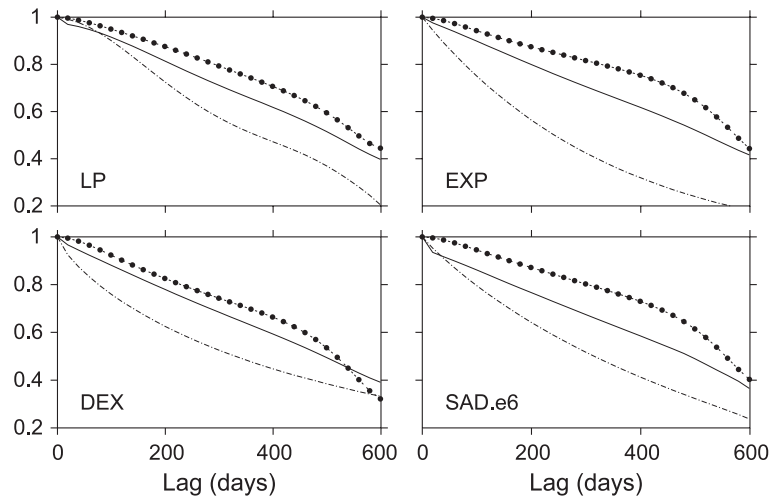


Fig. 4. Average estimates of phenotypic (solid line), direct genetic (●) and permanent environmental (dash-dot) correlations, according to the lag between the ages at recording, for models assuming $k_a=6$.

sistently high (not shown). As expected, differences between models occurred mainly for direct genetic and animal permanent environmental effect correlations, with those obtained by SAD closer to those estimated with a LP model.

4. Discussion

Independent of k_a , models assuming SAD with six or seven classes for residual variances had higher log L and smaller AIC and BIC than SAD with a cubic VF for residual variances. Presumably, this was due to the higher number of parameters available to model environmental variances not explained by animal permanent environmental effects.

Likelihood ratio tests (LRT) tend to favour models with high number of parameters, and information criteria such as AIC and BIC have been suggested instead to choose between models. However, variance estimates obtained with the best model (49 parameters) based on BIC ($k_a=6$, SAD and six classes for residual variances), were similar to those obtained with a model with 47 parameters ($k_a=6$, SAD and a cubic VF for residual variances), even with large differences in log L and BIC. This result suggests that the variance component estimates were fairly robust to changes in residual model as reported by Olori et al. (1999). On the other hand, when direct genetic effects

were modelled through a polynomial function of lower order ($k_a=4$), changes in the residual model had a huge impact on the other variance components estimates. It appears that the robustness of variance component estimates to changes in the residual would hold only if the other random effects were modelled with polynomials of sufficient order.

Another aspect to consider is that parameter estimates obtained by assuming a SAD correlation structure and $k_a=4$ were closer to those from standard univariate analysis than estimates from the best model on the basis of BIC ($k_a=6$, SAD and six classes for residual variances). Similarly, comparing a set of models for lactation curve analysis, Jaffrézic et al. (2002) found that the best model on the basis of likelihood criterion produced genetic variance estimates very different from the unstructured model. The authors concluded that likelihood may not be the most appropriate criterion to select a model.

Random regression models have become popular to deal with longitudinal traits and in the future may replace the usual multi-trait model used nowadays for genetic evaluation of repeated traits. Compared to multi-trait models, the random regression models will give more accurate predictions of breeding values (Meyer, 2004). Moreover, with RR, it is possible to predict breeding values for the growth curve as a whole and for functions of the growth curve. How-

ever, compared to multi-trait models, applying a random regression model will increase the number of mixed-model equations, the coefficient matrix of random effects will be denser, and computational requirements will increase accordingly. Previous analyses of the same data (Albuquerque and Meyer, 2001) suggested that a model with 77 parameters, chosen on the basis of AIC, was necessary to model the variances in the data. For BIC, with a more stringent penalty for the number of parameters, a model with 51 parameters fitted best. For both models, a Legendre polynomial of order 6 (quintic) was required to model animal permanent environmental effects. Modelling direct permanent environmental covariances through a nonstationary parametric correlation function combined with a variance function can decrease the number of parameters to be estimated. Results show that parameter estimates were similar to those obtained with a Legendre polynomial. Keeping the same model for the other random effects, it would be possible to decrease the number of parameters by 12–19 parameters. For lactation milk yield in dairy cattle, Jaffrézic et al. (2002) found that the genetic variance estimates obtained using a parametric function to model genetic and permanent environmental variance and correlation structures, were very different from those obtained with an unstructured model. However, using the same set of data, Jaffrézic et al. (2003), described results similar to the present work. The authors found that a third-order SAD model (6 parameters) fitted the environmental covariance structure better than a quartic polynomial (15 parameters). However, for the genetic part, a quadratic polynomial fitted best.

Although parameter estimates were similar, the change from a Legendre polynomial to a parametric correlation function was not without consequences. Differences in direct heritability between models applying parametric correlation functions and those applying Legendre polynomials were from -0.16 to 0.02 with averages from 0.02 to 0.04 when $k_a=6$ and from -0.06 to 0.09 with average of 0.04 when $k_a=4$. Future research should concentrate on quantifying differences in genetic parameter estimates and examine response to selection if a parametric correlation function was used to model animal permanent environmental covariances instead of Legendre polynomials.

5. Conclusion

Parametric correlation functions combined with variance functions to model animal permanent environmental covariances were compared to models fitting Legendre polynomial of age for this same effect. Results indicated that a nonstationary correlation function combined with a variance function modelled covariances for animal permanent environmental effects adequately. Furthermore, it can decrease the number of parameters to be estimated substantially, compared to a model fitting random regression on Legendre polynomial of age. A nonstationary function appears most appropriate. There are indications that the robustness of variance component estimates to changes in the model for residual variances depends on the order of polynomial fit for direct genetic effects. Further research is necessary to confirm these results and to evaluate the consequences of using parametric correlation functions to model animal permanent environmental covariances on response to selection.

Acknowledgement

This work was supported by FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) and by grant BFGEN.002 of Meat and Livestock Austrália (MLA). The authors thank ABCZ (Associação Brasileira dos Criadores de Zebu) for providing the data.

References

- Albuquerque, L.G., Meyer, K., 2001. Estimates of covariance functions for growth from birth to 630 days of age in Nelore cattle. *J. Anim. Sci.* 79, 2776–2789.
- Bermejo, J.L., Roche, R., Schulze, V., Rave, G., Looft, H., Kalm, E., 2003. Random regression to model genetically the longitudinal data of daily feed intake in growing pigs. *Livest. Prod. Sci.* 82, 189–199.
- Brotherstone, S., White, I.M.S., Meyer, K., 2000. Genetic modelling of daily milk yield using orthogonal polynomials and parametric curves. *Anim. Sci.* 70, 407–418.
- Carvalho, J., Pollak, E.J., Quaas, R.L., Blake, R.W., 2002. An autoregressive repeatability animal model for test-day records in multiple lactations. *J. Dairy Sci.* 85, 2040–2045.
- Diggle, P.J., Liang, K.Y., Zeger, S.L., 1994. *Analysis of Longitudinal Data*. Oxford Science Publications, Clarendon Press, Oxford.
- Foulley, J.-L., Jaffrézic, F., Robert-Granié, C., 2000. EM-REML estimation of covariance parameters in Gaussian mixed models for longitudinal data analysis. *Genet. Sel. Evol.* 32, 129–141.

- Gilmour, A.R., Cullis, B.R., Welham, S.J., Thompson, R., 1999. ASREML Reference Manual. NSW Agriculture Biometric Bulletin, vol. 3. NSW Agriculture, Australia.
- Harville, D.A., 1979. Recursive estimation using mixed linear models with autoregressive random effects. In: Van Vleck, L.D., Searle, S.R. (Eds.), Proceedings of a Conference in Honor of C. R. Henderson. Variance Components and Animal Breeding, pp. 157–179.
- Jaffrézic, F., Pletcher, S.D., 2000. Statistical models for estimating the genetic basis of repeated measures and other function-valued traits. *Genetics* 156, 913–922.
- Jaffrézic, F., White, I.M.S., Thompson, R., Visscher, P.M., 2002. Contrasting models for lactation curve analysis. *J. Dairy Sci.* 85, 968–975.
- Jaffrézic, F., White, I.M.S., Thompson, R., 2003. Use of score test as goodness-of-fit measure of the covariance structure in genetic analysis of longitudinal data. *Genet. Sel. Evol.* 35, 185–198.
- Kirkpatrick, M., Hill, W.G., Thompson, R., 1994. Estimating the covariance structure of traits during growth and ageing, illustrated with lactation in dairy cattle. *Genet. Res.* 64, 57–69.
- Littell, R.C., Pendersgast, J., Natarajan, R., 2000. Modelling covariance structure in the analysis of repeated measures data. *Statist. Med.* 19, 1793–1919.
- Mansour, H., Nordheim, E.V., Rutledge, J.J., 1985. Maximum likelihood estimation of variance components in repeated measures designs assuming autoregressive errors. *Biometrics* 41, 287–294.
- Meyer, K., 1998a. Estimating covariance functions for longitudinal data using a random regression model. *Genet. Sel. Evol.* 30, 221–240.
- Meyer, K., 1998b. “DXMRR”— a program to estimate covariance functions for longitudinal data by restricted maximum likelihood. *Proceedings 6th World Congr. Genet. Appl. Livest. Prod.*, Armidale, Australia, vol. 27, pp. 465–466.
- Meyer, K., 2001a. Estimates of direct and maternal covariance functions for growth of Australian beef calves from birth to weaning. *Genet. Sel. Evol.* 33, 1–28.
- Meyer, K., 2001b. Estimating genetic covariance functions assuming a parametric correlation structure for environmental effects. *Genet. Sel. Evol.* 33, 557–585.
- Meyer, K., 2004. Scope for a random regression model in genetic evaluation of beef cattle for growth. *Livest. Prod. Sci.* 86, 69–83.
- Nuñez-Anton, V., Zimmerman, D.L., 2000. Modeling nonstationary longitudinal data. *Biometrics* 56, 699–705.
- Olori, V.E., Hill, W.G., McGuirk, B.J., Brotherstone, S., 1999. Estimating variance components for test day milk records by restricted maximum likelihood with random regression animal model. *Livest. Prod. Sci.* 61, 53–63.
- Pletcher, S.D., Geyer, C.J., 1999. The genetic analysis of age-dependent traits: modeling the character process. *Genetics* 151, 825–835.
- SAS, 1997. SAS/STAT Software: Changes and Enhancements through Release 6.12. SAS Institute Inc., Cary.
- Wade, K.M., Quaas, R.L., Van Vleck, L.D., 1993. Estimation of the parameters involved in a first-order autoregressive process for contemporary groups. *J. Dairy Sci.* 76, 3033–3040.
- Wolfinger, R.D., 1993. Covariance structure in general mixed models. *Commun. Statist.* 22B, 1079–1106.