Original article

Estimating genetic covariance functions assuming a parametric correlation structure for environmental effects

Karin MEYER*

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

Abstract – A random regression model for the analysis of "repeated" records in animal breeding is described which combines a random regression approach for additive genetic and other random effects with the assumption of a parametric correlation structure for within animal covariances. Both stationary and non-stationary correlation models involving a small number of parameters are considered. Heterogeneity in within animal variances is modelled through polynomial variance functions. Estimation of parameters describing the dispersion structure of such model by restricted maximum likelihood *via* an "average information" algorithm is outlined. An application to mature weight records of beef cow is given, and results are contrasted to those from analyses fitting sets of random regression coefficients for permanent environmental effects.

repeated records / random regression model / correlation function / estimation / REML

1. INTRODUCTION

Random regression (RR) models have become a preferred choice in the analysis of longitudinal data in animal breeding applications. Typical applications have been the analysis of test day records in dairy cattle and growth or feed intake records in pigs and beef cattle; see, for instance, Meyer [25] for references.

RR models are particularly useful when we are interested in differences between individuals, as we obtain a complete description of the trajectory, *i.e.* "growth curve", over the range of ages considered. A popular model involves regression on (orthogonal) polynomials of time. This does not require prior assumptions about the shape of the trajectory. Such RR have proven to be well capable of modelling changes in variation due to distinct events, such

---

* Correspondence and reprints
E-mail: kmeyer@didgeridoo.une.edu.au
** A joint unit with NSW Agriculture

as weaning in beef cattle [25], or seasonal influences, *e.g.* [24]. However, frequently this required high orders of polynomial fit, and thus a large number of parameters to be estimated, accompanied by extensive computational requirements and numerical problems inherent to high order polynomials.

More generally in the analysis of longitudinal data, within-subject covariances between repeated records are often assumed to have a parametric correlation structure. In the simplest case, this might require a single parameter to specify correlations between records together with other parameters to model variances of records. A well-known example is the so-called "auto-correlation" structure. Other models involving a single parameter or low numbers of parameters (2, 3, 4) to model a correlation function are available, *e.g.* [2, 11, 29, 31, 40].

Pletcher and Geyer [33] presented an application of such models in the estimation of genetic covariance functions for age dependent traits in Drosophila. Their approach teamed a polynomial variance function (VF) to model changes in variances with age with a one-parameter correlation function (RF) to model correlations between different ages. However, estimation of the resulting covariance function (CF) used a reparameterisation of the covariance matrix among all ages in the data, as used by Meyer and Hill [27]. This resulted in computational requirements proportional to the number of ages in the data. Hence, their procedure is not readily applicable to large data sets arising from animal breeding applications with numerous different ages.

Recently, Foulley *et al.* [5] described an Expectation-Maximization type restricted maximum likelihood (REML) algorithm to estimate the covariance parameters for a model which combined a RR approach to model variation between subjects (*e.g.* genetic) with a single parameter RF to describe within subject covariances between repeated records. Their model included up to three parameters to model the latter, namely the parameter for the RF, the within subject variance and a measurement error variance.

Simple correlation models like those considered by Pletcher and Geyer [33] and Foulley *et al.* [5], generally imply stationarity, *i.e.* that the correlation between observations at any two times depends only on the difference between them, the "lag", not the times themselves. This might not be appropriate for animal breeding applications. Non-stationary correlation or covariance models are available, but usually involve more parameters. A common model, available in standard statistical analyses packages, is the so-called "ante-dependence" model [15]. A more parsimonious variant are structured ante-dependence models [32]. Pourahmadi [34] recently considered such models in a general mixed model framework.

This paper outlines REML estimation for RR models in animal breeding applications, assuming a parametric correlation structure for within animal covariances between repeated records. Both stationary and non-stationary

models are considered. A numerical example comprising the analysis of mature weight records of beef cows is presented.

2. MODEL OF ANALYSIS

2.1. Random regression model

RR models commonly applied in animal breeding include at least two sets of RR coefficients for each animal, representing direct, additive genetic and permanent environmental effects, respectively. Let y_{ij} denote the j-th record for animal i taken at time t_{ij}. Assume we fit RR on orthogonal polynomials of time or age at recording. The RR model is then

$$y_{ij} = F_{ij} + \sum_{m=0}^{k_A - 1} \alpha_{im} \phi_m(t_{ij}) + \sum_{m=0}^{k_R - 1} \gamma_{im} \phi_m(t_{ij}) + \varepsilon_{ij} \qquad (1)$$

with F_{ij} denoting the fixed effects pertaining to y_{ij} (often including a fixed regression on polynomials of time at recording), α_{im} and γ_{im} the additive genetic and permanent environmental RR coefficients for animal i, respectively, k_A and k_R the corresponding orders of polynomial fit, $\phi_m(t_{ij})$ the m-th orthogonal polynomial of time t_{ij} (standardised if applicable), and ε_{ij} the temporary environmental effect or "measurement error" affecting y_{ij}.

Let $\boldsymbol{\alpha}_i = \{\alpha_{im}\}$ and $\boldsymbol{\gamma}_i = \{\gamma_{im}\}$ denote the vectors of RR coefficients for animal i of length k_A and k_R, respectively. Assume a multivariate normal distribution of records y_{ij}, and

$$E[\boldsymbol{\alpha}_i] = \mathbf{0} \qquad\qquad E[\boldsymbol{\gamma}_i] = \mathbf{0}$$
$$Var(\boldsymbol{\alpha}_i) = \mathbf{K}_A \qquad Var(\boldsymbol{\gamma}_i) = \mathbf{K}_R$$
$$Cov(\boldsymbol{\alpha}_i, \boldsymbol{\gamma}_i') = \mathbf{0} \qquad Var(\boldsymbol{\varepsilon}_i) = Diag\{\sigma_{\varepsilon k}^2\}$$

with $\mathbf{K}_A = \{K_{A\,mn}\}$ and $\mathbf{K}_R = \{K_{R\,mn}\}$ the matrices of covariances among RR coefficients, and $\sigma_{\varepsilon k}^2$ the variances of measurement errors.

Further, let $\mathbf{y}_i$ be the ordered vector of observations for the i-th animal (ordered according to t_{ij}), and $\mathbf{y}$ of length M represent the complete vector of observations for all animals in the data, $i = 1, \ldots, N$. Assume relationships between animals are known and taken into account, incrementing the number of animals in the analysis through inclusion of parents without records to N_A. Let $\mathbf{b}$ of length N_F denote the vector of fixed effects to be fitted with design matrix $\mathbf{X}$, and $\boldsymbol{\alpha}$ of length $k_A \times N_A$ and $\boldsymbol{\gamma}$ of length $k_R \times N$ the vectors of additive genetic and permanent environmental RR coefficients. Design matrices for $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ have non-zero elements $\phi_m(t_{ij})$, *i.e.* orthogonal polynomials evaluated for the times at which measurements are recorded.

Let $\boldsymbol{\phi}$ of size $M \times k_R N$ denote the matrix of orthogonal polynomials evaluated for the ages in the data, with non-zero block of size $n_i \times k_R$ for the i-th animal. This is the design matrix for $\boldsymbol{\gamma}$. The corresponding matrix for $\boldsymbol{\alpha}$ is $\boldsymbol{\phi}_A$ of size $M \times k_A N_A$, augmented by columns of zero elements for animals without records. Finally, let $\boldsymbol{\varepsilon}$ denote the vector of measurement errors corresponding to $\mathbf{y}$. This gives

$$\mathbf{y} = \mathbf{Xb} + \boldsymbol{\phi}_A \boldsymbol{\alpha} + \boldsymbol{\phi} \boldsymbol{\gamma} + \boldsymbol{\varepsilon} \tag{2}$$

Let $\mathbf{y}$ and $\boldsymbol{\gamma}$ be ordered according to animals, $\boldsymbol{\alpha}$ be ordered according to RR coefficients. With $\mathbf{A}$ denoting the numerator relationship between animals and $\mathbf{I}_N$ an identity matrix of size N, this gives

$$\begin{aligned} Var(\mathbf{y}) &= \boldsymbol{\phi}_A \left(\mathbf{K}_A \otimes \mathbf{A} \right) \boldsymbol{\phi}_A' + \boldsymbol{\phi} \left(\mathbf{I}_N \otimes \mathbf{K}_R \right) \boldsymbol{\phi}' + Diag \left\{ \sigma_{\varepsilon k}^2 \right\} \\ &= \boldsymbol{\phi}_A \mathbf{G} \, \boldsymbol{\phi}_A' + \mathbf{R} + \boldsymbol{\Sigma}_\varepsilon = \mathbf{V} \end{aligned} \tag{3}$$

With measurement errors assumed uncorrelated, $\boldsymbol{\Sigma}_\varepsilon$ is diagonal and the mixed model equations and matrix (MMM) pertaining to (2) can be set up as for univariate analyses. Moreover, if $\boldsymbol{\Sigma}_\varepsilon = \sigma_\varepsilon^2 \mathbf{I}_M$ or $Diag \left\{ \sigma_\varepsilon^2 d_k \right\}$, σ_ε^2 can be factored from the MMM and be estimated directly from the residual sum of squares [22].

The covariance function due to permanent environmental effects of the animal ($\mathcal{R}$) is estimated through $\mathbf{K}_R$. With $\mathcal{R}$ generally fitted to reduced order, *i.e.* k_R smaller than the number of ages in the data, the resulting estimate of $\mathbf{R}$, the permanent environmental covariance matrix among observations, is smoothed and has reduced rank. However, it does not have a pre-imposed structure. Whilst it is straightforward to estimate $\mathbf{K}_R$ assuming a certain structure, this does not translate readily to $\mathbf{R}$.

Equivalent model

We are, however, more interested in imposing a structure on $\mathbf{R}$ than $\mathbf{K}_R$. This can be achieved by fitting an equivalent model to (2)

$$\mathbf{y} = \mathbf{Xb} + \boldsymbol{\phi}_A \boldsymbol{\alpha} + \mathbf{e} \tag{4}$$

with $\mathbf{e}$ of length M the vector of total environmental effects, *i.e.* the sum of permanent effects due to the animal and measurement errors. This has variance

$$Var(\mathbf{e}) = \mathbf{R}^* = \mathbf{R} + \boldsymbol{\Sigma}_\varepsilon \tag{5}$$

Like $\mathbf{R}$, $\mathbf{R}^*$ is blockdiagonal for animals. Permanent environmental covariances between records taken on the same animal are modelled through non-zero off-diagonal elements in the i-th block of $\mathbf{R}^*$, $\mathbf{R}_i^*$. The MMM for (5) can be set up for one animal at a time, as for standard, non-RR multivariate analyses.

They can be thought of as derived from the MMM for (2) by absorbing $\boldsymbol{\gamma}$, and computational requirements to factor the MMM are the same for both models.

Choosing (5) rather than (2), however, offers a much wider choice of parameterisation for $\mathbf{R}^*$ and $\mathbf{R}$, and allows for a chosen structure of $\mathbf{R}$ to be imposed easily.

2.2. Parametric correlation structures

Decompose $\mathbf{R}$ into the product of standard deviations and correlations

$$\mathbf{R} = \boldsymbol{\Sigma}_R^{1/2} \, \mathbf{C} \, \boldsymbol{\Sigma}_R^{1/2} \tag{6}$$

with $\boldsymbol{\Sigma}_R = Diag\left\{\sigma_{Rj}^2\right\}$ the diagonal matrix of permanent environmental variances pertaining to $\mathbf{y}$, and $\mathbf{C} = \left\{c_{jk}\right\}$ the corresponding matrix of correlations. $\mathbf{C}$ is blockdiagonal for animals.

2.2.1. Variance function

Heterogeneous variances have been modelled through VF, *e.g.* [6,33,34], and this has been applied to measurement error variances in RR analyses [12, 25,35]. Similarly, we can model the *j*-th element of $\boldsymbol{\Sigma}_R$ or $\boldsymbol{\Sigma}_R^{1/2}$ as a function of the age at recording t_{ij}. This can be a step function or, as more commonly used, a polynomial function, For instance,

$$\sigma_{Rj}^w = \sigma_{R0}^w \left(1 + \sum_{r=1}^{v} \beta_r \, t_{ij}^r\right) \tag{7}$$

$$\sigma_{Rj}^w = \sigma_{R0}^w \, e^{\left(1 + \sum_{r=1}^{v} \beta_r \, t_{ij}^r\right)} \tag{8}$$

$$\sigma_{Rj}^w = e^{\left(\sigma_{R0}^w + \sum_{r=1}^{v} \beta_r \, t_{ij}^r\right)} \quad \text{or} \quad \log\left(\sigma_{Rj}^w\right) = \sigma_{R0}^w + \sum_{r=1}^{v} \beta_r \, t_{ij}^r \tag{9}$$

with σ_{R0}^2 the variance at the intercept, β_r the coefficients of the VF and v the order of polynomial fit. Either variances ($w = 2$) or standard deviations ($w = 1$) can be modelled in this way. Functions (8) and (9) are advantageous when variances increase exponentially with time. In addition, they require less restrictions on the parameters of the VF than (7) to ensure that $\sigma_{Rj}^2 > 0$ for all $j = 1, \ldots, M$. Whilst functions shown above involve ordinary polynomials (as in previous applications), use of orthogonal polynomials of t_{ij} may be preferable to reduce sampling correlations between β_r and thus improve convergence when estimating these parameters. Alternatively, for applications where variances show some periodicity, *e.g.* due to seasonal influences, a VF involving both polynomial and trigonometric terms [4] may be beneficial. In other instances, segmented polynomials [7] may be able to model changes in variances with time with fewer parameters.

2.2.2. Correlation function

Correlations between observations at different ages can be modelled as a function of the ages and one or more parameters of the RF. Correlation functions are *stationary* if a correlation between a pair of records depends only on the differences in ages at which they were taken –or lag– rather than the ages themselves. Most popular RF, including those given by (11) to (19) below, fall into this category.

Compound symmetry

In the simplest case, correlations between all observations for an animal (at different ages) are assumed to be the same.

$$c_{jk} = \begin{cases} 1 & \text{for } j = k \\ \rho & \text{for } j \neq k \end{cases} \tag{10}$$

with ρ a correlation, *i.e.* $-1 < \rho < 1$. This pattern is generally referred to as uniform correlation or compound symmetry (CS), and is the correlation structure assumed in the standard "repeatability model" analyses often used in the analysis of animal breeding data.

Auto-correlation

Let $\ell_{jk} = |t_{ij} - t_{ik}|$ denote the lag in ages for a pair of records (y_{ij}, y_{ik}) on the i-th animal. The so-called power, serial or auto-correlation function is then

$$c_{jk} = \rho^{\ell_{jk}} \tag{11}$$

with $-1 < \rho < 1$ as above. This is the correlation structure generated by a continuous-time, first order auto-regressive (AR(1)) process.

Exponential model

An alternative way to model the correlation structure given by (11) is the exponential (EXP) model

$$c_{jk} = e^{-\theta \ell_{jk}} \tag{12}$$

with $\theta = -\log(\rho) > 0$. Again this parameterisation can be advantageous in terms of estimation, as it does not require the parameter of the RF to be constrained to an interval.

Gaussian model

In some instances, the decline in correlation with increasing lag is steeper than can be modelled with an exponential function of ℓ_{jk}. In this case, the so-called Gaussian (GAU) exponential model which uses ℓ_{jk}^2 may be more appropriate.

$$c_{jk} = e^{-\theta \, \ell_{jk}^2} \tag{13}$$

Diggle *et al.* [3] emphasize that in contrast to EXP, GAU is differentiable at $\ell_{jk} = 0$, and that for a sufficiently small time scale GAU has smoother appearance than EXP.

Other single parameter functions

Other RF involving different distributions but only a single parameter have been examined by Pletcher and Geyer [33]. All yield correlations which decrease with increasing lag. For instance,

$$c_{jk} = \left(1 + \theta \ell_{jk}^2\right)^{-1} \tag{14}$$

$$c_{jk} = \left(\cosh(\pi \theta \ell_{jk}/2)\right)^{-1} \tag{15}$$

$$c_{jk} = \sin(\theta \ell_{jk}) / \left(\theta \ell_{jk}\right) \tag{16}$$

$$c_{jk} = \left(1 - \cos(\theta \ell_{jk})\right) / \left(\theta^2 \ell_{jk}^2\right) \tag{17}$$

are RFs based on the Cauchy distribution, the hyperbolic cosine, the characteristic function of the uniform distribution, and the characteristic function of the triangular distribution, respectively.

"Damped" exponential model

A more flexible model can be obtained by adding a second parameter κ. This is a scale parameter which allows the exponential decay of the auto-correlation function to be accelerated or attenuated. Muñoz *et al.* [29] presented this for the serial correlation model

$$c_{jk} = \rho^{\ell_{jk}^\kappa} \tag{18}$$

pointing out that for $\kappa = 1$, $\kappa = 0$ and $\kappa = \infty$, (18) reduces to the serial correlation, compound symmetry and first-order moving average model, respectively. Alternatively, (12) can be expanded to

$$c_{jk} = e^{-\theta \, \ell_{jk}^\kappa} \tag{19}$$

[11]. Pletcher and Guyer [33] consider this as RF based on the characteristic function of the general stable distribution, with restriction $0 < \kappa < 2$. In the following, (19) is referred to as damped exponential (DEX) model.

Other two parameter functions

A model which is not a special case of DEX, is the RF generated by a second-order auto-regressive process, which has parameters determined by the correlation between ages with lags 1 and 2 [29].

Any of the above RF ((11) to (19)) can be modified to allow for a proportion τ of the correlation independent of age effects [11]

$$c_{jk} = \tau + (1 - \tau)\, c_{jk}^* \tag{20}$$

with c_{jk}^* a function of the lag in ages as modelled above, and τ estimated as an additional parameter (yielding a three-parameter RF if extending (19)).

Structured ante-dependence model

Another class of models employed in the analysis of "repeated" records or longitudinal data are the so-called *ante-dependence* (AD) models, *e.g.* [15]. These are loosely related to time series models, in such that the *j*-th record on an animal depends on and is correlated to a number of its predecessors [3]. In contrast to the parametric correlation structures considered so far, AD models allow for non-stationary correlations.

For an AD model of order *s*, AD(*s*), a record y_{ij} in the ordered vector of observations $\mathbf{y}_i$ for animal *i* is assumed to depend at most on records $y_{i(j-1)}, \ldots, y_{i(j-s)}$, but to be independent of any other preceding observations $y_{i(j-s-1)}, \ldots, y_{i1}$. This yields a correlation matrix with elements on the first *s* subdiagonals as variables, and the elements of the remaining subdiagonals $(s+1, \ldots, n-1)$ determined by the former. Consequently, the corresponding inverse is a banded matrix, with only the elements of the leading diagonal and first *s* subdiagonals being non-zero [15]. Hence, for *n* different times of recording, an unstructured AD(*s*) model has $(s+1)(2n-s)/2$ parameters, *n* variances and $sn - s(s+1)/2$ correlations.

For $s = 1$, a first-order AD model, there are $n-1$ correlations on the first sub-diagonal of the correlation matrix, $c_{j(j+1)}$. The other correlations are given by a simple multiplicative relation

$$c_{jk} = \prod_{q=j+1}^{k-1} c_{q(q+1)} \qquad \text{for } j = 1, n \text{ and } k = j+2, n \tag{21}$$

[31]. For $s > 1$, the functional relationship with the elements of the first *s* subdiagonals is more complicated. In that case, a parameterisation in terms of the inverse of the corresponding covariance matrix – also called the "concentration matrix" of the AD – or it's Cholesky decomposition is often preferred.

Whilst an AD(*s*) model with low *s* has considerably less parameters than a full multivariate, unstructured model (which has $n(n+1)/2$ parameters), it

can still involve impractically many parameters. *Structured* ante-dependence (SAD) models [32] assume a functional relationship between the parameters of an AD model, and thus provide a more parsimonious representation. Firstly, variances are considered to be a function of the time at measurement, with the function involving a small number of parameters. Zimmerman *et al.* [39], Núñez-Antón and Zimmerman [32] and Pourahmadi [34] consider polynomial VFs as described above (see (7) to (9)). Secondly, the correlations on the first s subdiagonals are determined by the times of recording and $2s$ parameters, ρ_k and κ_k, respectively:

$$c_{j\,(j-k)} = \rho_k^{f(t_{ij},\,\kappa_k)\,-\,f(t_{i\,(j-k)},\,\kappa_k)} \qquad \text{for } k = 1, s \text{ and } j = k+1, n \quad (22)$$

[32] with $0 < \rho < 1$ and

$$f(t_{ij}, \kappa_k) = \begin{cases} \left(t_{ij}^{\kappa_k} - 1\right)/\kappa_k & \text{for } \kappa_k \neq 0 \\ \log(t_{ij}) & \text{for } \kappa_k = 0 \end{cases} \qquad (23)$$

Function (23) applies a deformation (Box-Cox power transformation) to the time scale which facilitates non-stationarity of correlations. For $\kappa < 1$ equidistant correlations are increasing with age. Conversely, $\kappa > 1$ implies lower correlations between records with equal lag at higher ages [39]. For $s = 1$ and $\kappa = 1$, the RF (22) reduces to (11), the auto-correlation function.

3. ESTIMATION OF COVARIANCE AND CORRELATION FUNCTIONS

Parameters of covariance, correlation and variance functions are readily estimated by restricted maximum likelihood (REML). This may involve a derivative-free procedure, an "Average Information" (AI-REML) algorithm [8] or an Expectation-Maximization (EM) algorithm, as described by Foulley *et al.* [5]. Various authors consider REML estimation in the analysis of longitudinal or spatial data, but often do not go further than specifying the log likelihood and using a simple search procedure, such as the simplex method of Nelder and Mead [30], to locate its maximum, *e.g.* [3,31,39]. Others describe maximum likelihood estimation using Newton-Raphson type algorithms, *e.g.* [13,18,29]. Gilmour *et al.* [8] consider AI-REML estimation for models with correlated residuals in a general formulation.

3.1. The likelihood

The REML log likelihood for (4) is

$$-2\log \mathcal{L} = const + \log |\mathbf{G}| + \log \left|\mathbf{R}^*\right| + \log |\mathbf{C}_M| + \mathbf{y}'\mathbf{P}\mathbf{y} \qquad (24)$$

where $\mathbf{C}_M$ is the coefficient matrix in the mixed model equation pertaining to (4) and $\mathbf{y'Py}$ is the sum of squares of residuals. Both $\mathbf{y'Py}$ and $\log|\mathbf{C}_M|$ can be evaluated simultaneously as described by Graser *et al.* [9], by factoring the corresponding MMM

$$\mathbf{M} = \begin{bmatrix} \mathbf{X'(R^*)}^{-1}\mathbf{X} & \mathbf{X'(R^*)}^{-1}\boldsymbol{\phi}_A & \mathbf{X'(R^*)}^{-1}\mathbf{y} \\ \boldsymbol{\phi}'_A\mathbf{(R^*)}^{-1}\mathbf{X} & \boldsymbol{\phi}'_A\mathbf{(R^*)}^{-1}\boldsymbol{\phi}_A + \mathbf{K}_A^{-1} \otimes \mathbf{A}^{-1} & \boldsymbol{\phi}'_A\mathbf{(R^*)}^{-1}\mathbf{y} \\ \mathbf{y'(R^*)}^{-1}\mathbf{X} & \mathbf{y'(R^*)}^{-1}\boldsymbol{\phi}_A & \mathbf{y'(R^*)}^{-1}\mathbf{y} \end{bmatrix} \quad (25)$$

$\mathbf{M}$ is large but sparse, with $N_M = N_F + k_A N_A + 1$ rows and columns. For $\mathbf{R}^*$ blockdiagonal it can be set up for one animal at a time, as for corresponding multivariate analyses. Factoring $\mathbf{M}$ into $\mathbf{LL'}$ with $\mathbf{L}$ a lower triangular matrix with elements l_{ij} ($l_{ij} = 0$ for $j > i$) gives

$$\log|\mathbf{C}_M| = 2 \sum_{k=1}^{N_M-1} \log l_{kk} \qquad \text{and} \qquad (26)$$

$$\mathbf{y'Py} = l^2_{N_M N_M} \qquad (27)$$

e.g. [28]. The other components of (24) can be evaluated as

$$\log|\mathbf{G}| = N_A \log|\mathbf{K}_A| + k_A \log|\mathbf{A}| \qquad \text{and} \qquad (28)$$

$$\log|\mathbf{R}^*| = \sum_{i=1}^{N} \log|\mathbf{R}_i + \Sigma_{\varepsilon i}| \qquad (29)$$

This involves determinants of small matrices only, of size k_A and the number of records for each animal, respectively. For some correlation structures, closed forms for the corresponding inverse correlation or covariance matrices and determinants exist. In some cases, in particular for analyses assuming $\Sigma_\varepsilon = \mathbf{0}$, this can be exploited to reduce computational requirements to evaluate (29).

3.2. AI-REML algorithm

Maximisation of $\log \mathcal{L}$ *via* AI-REML requires first derivatives of (24) and the average of observed and expected information [8]. The latter is proportional to second derivatives of the data part, $\mathbf{y'Py}$, of the likelihood. These can be determined as for standard multivariate analyses, using sparse matrix inversion and repeated solution of the mixed model equations [19] or automatic differentiation of the MMM [20].

3.2.1. First derivatives

Derivatives of $\log|\mathbf{C}_M|$ and $\mathbf{y}'\mathbf{P}\mathbf{y}$ can be determined through automatic differentiation of the Cholesky factor of $\mathbf{M}$, as described by Smith [36]. This requires the derivatives of $\mathbf{M}$ with respect to the parameters to be estimated.

$$\frac{\partial \mathbf{M}}{\partial K_{A\,mn}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{K}_A^{-1}\dfrac{\partial \mathbf{K}_A}{\partial K_{A\,mn}}\mathbf{K}_A^{-1} \otimes \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \tag{30}$$

for covariances among the RR coefficients for additive genetic effects, α_i. The derivative $\partial \mathbf{K}_A / \partial K_{A\,mn}$ has elements of unity in position mn and nm and zero otherwise. Hence,

$$-\mathbf{K}_A^{-1}\frac{\partial \mathbf{K}_A}{\partial K_{A\,mn}}\mathbf{K}_A^{-1} = \left\{ -\frac{2-\delta_{mn}}{2}\left(K_A^{rm}K_A^{ns} + K_A^{rn}K_A^{ms}\right)\right\}$$
$$\text{for } r,s = 1,\ldots,k_A \quad (31)$$

with K_A^{mn} the mn-th element of $\mathbf{K}_A^{-1}$ and δ_{mn} Kronecker's delta, i.e. $\delta_{mn} = 1$ for $m = n$ and 0 otherwise. Similarly,

$$\frac{\partial \mathbf{M}}{\partial \theta_R} = \begin{bmatrix} \mathbf{X}'\mathbf{Q}_R\mathbf{X} & \mathbf{X}'\mathbf{Q}_R\boldsymbol{\phi}_A & \mathbf{X}'\mathbf{Q}_R\mathbf{y} \\ \boldsymbol{\phi}_A'\mathbf{Q}_R\mathbf{X} & \boldsymbol{\phi}_A'\mathbf{Q}_R\boldsymbol{\phi}_A & \boldsymbol{\phi}_A'\mathbf{Q}_R\mathbf{y} \\ \mathbf{y}'\mathbf{Q}_R\mathbf{X} & \mathbf{y}'\mathbf{Q}_R\boldsymbol{\phi}_A & \mathbf{y}'\mathbf{Q}_R\mathbf{y} \end{bmatrix} \tag{32}$$

with

$$\mathbf{Q}_R = -(\mathbf{R}^*)^{-1}\frac{\partial \mathbf{R}^*}{\partial \theta_R}(\mathbf{R}^*)^{-1}$$

and θ_R standing in turn for the parameters of the permanent environmental VF ($\mathscr{P}$), σ_{R0}^2 and regression coefficient β_r, the parameters of the CF, ρ or θ and κ, and the parameters of the VF for measurement error variances ($\mathscr{E}$).

Let $\partial \mathbf{L}/\partial \theta_R$, with elements $\partial l_{ij}/\partial \theta_R$, denote the derivative of $\mathbf{L}$ with respect to θ_R obtained by differentiation of $\mathbf{L}$ as described by Smith [36]. First derivatives of the last two terms in (24) are then [28]

$$\frac{\partial \log|\mathbf{C}_M|}{\partial \theta_R} = 2\sum_{k=1}^{N_M-1} l_{kk}^{-1}\frac{\partial l_{kk}}{\partial \theta_R} \qquad \text{and} \tag{33}$$

$$\frac{\partial \mathbf{y}'\mathbf{P}\mathbf{y}}{\partial \theta_R} = 2\,l_{N_M N_M}\frac{\partial l_{N_M N_M}}{\partial \theta_R} \tag{34}$$

Derivatives of the other two terms in (24) can be evaluated indirectly

$$\frac{\partial \log |\mathbf{G}|}{\partial K_{A\,mn}} = N_A \, tr\left(\mathbf{K}_A^{-1}\frac{\partial \mathbf{K}_A}{\partial K_{A\,mn}}\right) = (2 - \delta_{mn})\, N_A \, K_A^{mn} \qquad (35)$$

$$\frac{\partial \log \left|\mathbf{R}^*\right|}{\partial \theta_R} = \sum_{i=1}^{N} tr\left((\mathbf{R}_i^*)^{-1}\frac{\partial \mathbf{R}_i^*}{\partial \theta_R}\right) \qquad (36)$$

For $\mathbf{R}^*$ linear in the parameters to be estimated, (36) simplifies analogously to (35), see [28].

3.2.2. Second derivatives

AI-REML algorithms [8,14,19,20] have generally considered the case where $\mathbf{V}$ is linear in the parameters to be estimated, as, for instance, for standard multivariate analyses. This gives second derivatives of $\mathbf{V}$ which are zero, and the average of "observed" and "expected" information for parameters θ_r and θ_s is equal to

$$-\frac{1}{2}\mathbf{y}'\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_r}\mathbf{P}\frac{\partial \mathbf{V}}{\partial \theta_s}\mathbf{P}\mathbf{y} \qquad (37)$$

For cases where $\mathbf{V} \neq \sum_p \theta_p(\partial \mathbf{V}/\partial \theta_p)$, as for our parameterisation of the residual covariance matrix in terms of a variance and parametric correlation function, second derivatives of $\mathbf{V}$ are non-zero. For such models, Gilmour *et al.* [8] suggest to approximate the exact average by a "simplified average" information. This is derived by approximating $\partial^2 \mathbf{y}'\mathbf{P}\mathbf{y}/\partial \theta_r \partial \theta_s$ by its expectation. Asymptotically the two are the same. Computationally, this is equivalent to ignoring extra terms involving non-zero second derivatives of $\mathbf{V}$.

Rewrite (37) as $\mathbf{b}_r'\mathbf{P}\mathbf{b}_s$ with $\mathbf{b}_r = \partial \mathbf{V}/\partial \theta_r \mathbf{P}\mathbf{y}$. For $\theta_r = K_{A\,mn}$,

$$\mathbf{b}_r = \boldsymbol{\phi}_A\left(\frac{\partial \mathbf{K}_A}{\partial k_{A\,mn}}\mathbf{K}_A^{-1}\otimes \mathbf{I}_{N_A}\right)\hat{\boldsymbol{\alpha}} = \frac{2 - \delta_{mn}}{2}\sum_{q=1}^{k_A}\left(K_A^{mq}\boldsymbol{\phi}_{A\,n} + K_A^{nq}\boldsymbol{\phi}_{A\,m}\right)\hat{\boldsymbol{\alpha}}_q \quad (38)$$

with $\boldsymbol{\phi}_{A\,q}$ and $\boldsymbol{\alpha}_q$ denoting the sub-matrix of $\boldsymbol{\phi}_A$ and subvector of $\hat{\boldsymbol{\alpha}}$, respectively, for the q-th RR coefficient. Similarly, for θ_r a parameter of $\mathbf{R}^*$ and $\hat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\hat{\mathbf{b}} - \boldsymbol{\phi}_A\hat{\boldsymbol{\alpha}}$,

$$\mathbf{b}_r = \left(\sum_{i=1}^{N} {}^{+}\frac{\partial \mathbf{R}_i^*}{\partial \theta_r}(\mathbf{R}_i^*)^{-1}\right)\hat{\mathbf{e}} \qquad (39)$$

where "+" denotes the direct matrix sum. As shown previously [20], crossproducts $\mathbf{b}_r'\mathbf{P}\mathbf{b}_s$ can be evaluated by replacing $\mathbf{y}$ in (25) with $\mathbf{B} =$

$[\mathbf{b}_1|\mathbf{b}_2|\ldots|\mathbf{b}_p]$, a matrix containing one column $\mathbf{b}_i$ for each parameter, and expanding the MMM to

$$\mathbf{M}_B = \begin{bmatrix} \mathbf{C}_M & \mathbf{X}'(\mathbf{R}^*)^{-1}\mathbf{B} \\ & \boldsymbol{\phi}'_A(\mathbf{R}^*)^{-1}\mathbf{B} \\ \mathbf{B}'(\mathbf{R}^*)^{-1}\mathbf{X} \;\; \mathbf{B}'(\mathbf{R}^*)^{-1}\boldsymbol{\phi}_A & \mathbf{B}'(\mathbf{R}^*)^{-1}\mathbf{B} \end{bmatrix} \tag{40}$$

Factoring $\mathbf{M}_B$ then overwrites $\mathbf{B}'\mathbf{PB}$ with elements $\mathbf{b}'_r\mathbf{Pb}_s$. With the Cholesky factor of $\mathbf{C}_M$ already evaluated (in factoring $\mathbf{M}$), this is computationally undemanding.

3.2.3. Derivatives of $\mathbf{R}^*$

Evaluation of $\mathbf{Q}_R$ (33), the first derivatives of $\mathbf{R}^*$ (36) and vectors $\mathbf{b}_r$ pertaining to parameters of $\mathbf{R}^*$ (39) requires the partial derivatives of $\mathbf{R}^*$ with respect to the parameters to be estimated, as well as products and traces involving the inverse of $\mathbf{R}^*$. Corresponding terms for the parameters of a polynomial VF for measurement error variances under model (2), *i.e.* the simple case of a diagonal residual covariance matrix, have been given by Meyer [25].

Correlation function

For $\theta_R = \rho, \theta$ or κ

$$\frac{\partial \mathbf{R}^*_i}{\partial \theta_R} = \Sigma_R^{1/2} \frac{\partial \mathbf{C}_i}{\partial \theta_R} \Sigma_R^{1/2} = \left\{ \sigma_{Rj} \frac{\partial c_{jk}}{\partial \theta_R} \sigma_{Rk} \right\} \tag{41}$$

Derivatives of c_{jk} need to be determined for each RF separately. For instance, for the auto-correlation function, (11) or (18)

$$\frac{\partial c_{jk}}{\partial \rho} = \ell_{jk}^{\kappa} \rho^{\ell_{jk}^{\kappa} - 1} \tag{42}$$

$$\frac{\partial c_{jk}}{\partial \kappa} = \rho^{\ell_{jk}^{\kappa}} \ell_{jk}^{\kappa} \ln(\rho) \ln(\ell_{jk}) \tag{43}$$

Similarly, for an exponential model, EXP (12), GAU (13) or DEX (19),

$$\frac{\partial c_{jk}}{\partial \theta} = -\ell_{jk}^{\kappa} e^{-\theta \ell_{jk}^{\kappa}} \tag{44}$$

$$\frac{\partial c_{jk}}{\partial \kappa} = -\theta \ell_{jk}^{\kappa} \ln(\ell_{jk}) e^{-\theta \ell_{jk}^{\kappa}} \tag{45}$$

with fixed values for κ where appropriate.

Table I. Derivatives of permanent environmental standard deviations with respect to parameters of functions modelling changes in variances or standard deviations with time.

Function [a]	Standard Deviation		Variance	
	$\partial\sigma_{Rj}/\partial\sigma_{R0}$	$\partial\sigma_{Rj}/\partial\beta_r$	$\partial\sigma_{Rj}/\partial\sigma_{R0}^2$	$\partial\sigma_{Rj}/\partial\beta_r$
$\sigma_{Rj}^w = \sigma_{R0}^w \left(1 + \sum_{r=1}^v \beta_r t_{ij}^r\right)$	σ_{Rj}/σ_{R0}	$\sigma_{R0} t_{ij}^r$	$\frac{1}{2}\sigma_{Rj}/\sigma_{R0}^2$	$\frac{1}{2}\sigma_{Rj}^{-1} \sigma_{R0}^2 t_{ij}^r$
$\sigma_{Rj}^w = \sigma_{R0}^w \exp\left\{1 + \sum_{r=1}^v \beta_r t_{ij}^r\right\}$	σ_{Rj}/σ_{R0}	$\sigma_{Rj} t_{ij}^r$	$\frac{1}{2}\sigma_{Rj}/\sigma_{R0}^2$	$\frac{1}{2}\sigma_{Rj} t_{ij}^r$
$\sigma_{Rj}^w = \exp\left\{\sigma_{R0}^w + \sum_{r=1}^v \beta_r t_{ij}^r\right\}$	σ_{Rj}	$\sigma_{Rj} t_{ij}^r$	$\frac{1}{2}\sigma_{Rj}$	$\frac{1}{2}\sigma_{Rj} t_{ij}^r$

[a] $w = 1$ to model standard deviations, $w = 2$ to model variances.

Variance functions

Derivatives of the diagonal matrix of measurement error variances are

$$\frac{\partial \boldsymbol{\Sigma}_\varepsilon}{\partial \theta_\varepsilon} = Diag\left\{\frac{\partial \sigma_{\varepsilon j}^2}{\partial \theta_\varepsilon}\right\} \tag{46}$$

with θ_ε standing for the parameters modelling changes in temporary environmental variances over time. Expressions for parameters of a VF, $\mathcal{E}$, are analogous to those given below for permanent environmental variances.

For the parameters of $\mathcal{P}$, $\theta_R = \sigma_{R0}^2$ or σ_{R0} and β_r for $r = 1, \ldots, v$,

$$\frac{\partial \mathbf{R}^*}{\partial \theta_R} = \left\{\frac{\partial \sigma_{Rj}}{\partial \theta_R} c_{jk} \sigma_{Rk} + \sigma_{Rj} c_{jk} \frac{\partial \sigma_{Rk}}{\partial \theta_R}\right\} \tag{47}$$

Derivatives of σ_{Rj} depend on the function and parameterisation chosen. Values of $\partial\sigma_{Rj}/\partial\theta_R$ for functions (7) to (9) to model either standard deviations or variances are summarised in Table I.

4. APPLICATION

4.1. Data

Data consisted of January weights of Polled Hereford beef cows between 2 and 10 years of age, analysed previously fitting RR on Legendre polynomials of age for both additive genetic and permanent environmental effects of the animal [23].

Records originated from a selection experiment carried out at the Wokalup research station in the South West of Western Australia; see [26] for details. With a Mediterranean climate, characterised by Winter and Spring rains and pasture growth, and subsequent almost complete drought in Summer and Autumn, January weighings tended to record cows at their top weight during the year. Short mating periods resulted in the bulk of calves born in April and May each year. Ages at weighing for records selected thus ranged from 19 to 119 months, with up to 9 weights per cow. In total there were 3 320 records on 850 cows, offspring of 149 sires and 423 dams, with 165, 153, 103, 124, 83, 75, 47, 50 and 50 animals having 1, 2, ..., 8 and 9 records available.

4.2. Analyses

Data were analysed assuming a parametric correlation structure for covariances between records taken on the same animal. Models CS, EXP, GAU, DEX and SAD were considered, teamed with polynomial functions to model permanent environmental standard deviations ((7) with $w = 1$) of order $v = 0$ to 7. Measurement error variances were in turn considered homogeneous ($e = 1$), to be represented by a polynomial function ((7) with $w = 1$) as for permanent environmental effects, or to change with year of age, fitting a step function with $e = 7$ classes (2, 3, 4, 5, 6, 7, 8–10 years) as in previous analyses [23]. In the following "RF.VFv.Ee" denotes an analysis fitting correlation function RF with polynomial function $\mathcal{P}$ for permanent environmental standard deviations to order v, and e measurement error variances. A model fitting a polynomial function $\mathcal{E}$ for temporary environmental standard deviations with v' coefficients is described as "RF.VFv.E1+VFv'". For example, "SAD.VF3.E1" represents an analysis assuming a structured ante-dependence model for the correlations between records for an animal with a cubic variance function to model changes in permanent environmental variances over time, and temporary environmental effects which are considered to have homogeneous variances.

For comparison, data were also reanalysed assuming permanent environmental covariances followed a pattern as described by fitting a set of RR coefficients on Legendre polynomials (LP) of age at recording, *i.e.* $\mathbf{R} = \boldsymbol{\phi}\mathbf{K}_R\boldsymbol{\phi}'$. As in previous analyses, estimates of $\mathbf{K}_R$ were forced to have reduced rank, by setting eigenvalues less than 0.001 to zero, thus reducing the number of parameters to be estimated accordingly. In the following, "LP.Rk_Rrk_R'.Ee(+VFv')" denotes an analysis fitting LPs for permanent environmental effects to order k_R, with estimated covariance matrix of rank k_R' (other terms as above).

Analyses were first carried out on a "phenotypic" level, assuming all covariances between records were due to animals' permanent environmental effects. This facilitated computationally undemanding examination of a wide range of variance functions and correlation structures. Secondly, analyses allowed for (co)variances between individuals by fitting RR coefficients on LPs of age

due to animals' additive genetic effects, incorporating all pedigree information available. "Ak_Ark'_A" denotes an analysis fitting LPs to order k_A with estimated covariance matrix $\mathbf{K}_A$ of rank k'_A.

As in previous analyses, fixed effects fitted included contemporary groups, defined as year-paddock subclasses, and a cubic regression on Legendre polynomials of age. Parameters of variance, correlation and covariance functions were estimated by REML, using a combination of derivative-free and AI-REML algorithms. Analyses were carried out using program DxMRR[1] [21], extended to accommodate model (4) with parameterisation of $\mathbf{R}^*$ as described above.

Results from different analyses were compared by examining estimated variances and correlations for ages represented in the data. In addition, maximum values of log $\mathcal{L}$ and the REML forms of Akaike's (AIC) and Schwarz' (BIC) information criterion, *e.g.* [38] were contrasted.

4.3. Results

4.3.1. "Phenotypic" analyses

Results from analyses ignoring variation between animals are summarised in Table II. Values of log $\mathcal{L}$ are clearly dominated by the order of fit for VF $\mathcal{P}$ and the degree of heterogeneity allowed for measurement error variances. For $e = 1$, $\mathcal{P}$ needed to be a cubic or higher order polynomial, for estimates of variances at mean age not to be drastic over- or underestimates. Estimated correlations between records one month apart were close to unity. At equal order of fit for $\mathcal{P}$, v, two-parameter RFs (DEX and SAD) yielded higher log $\mathcal{L}$ than single parameter RFs (CS, EXP and GAU), but there was no advantage of the non-stationary SAD over the stationary damped exponential correlation structure.

Assuming homogeneous σ_ε^2, parametric RF had higher log $\mathcal{L}$ than analyses fitting LP involving similar numbers of parameters, presumably because more parameters were available to model changes in variation with time. For $e = 7$, however, there was little difference, with the heterogeneous σ_ε^2 accounting for differences in variation not modelled by covariance function $\mathcal{R}$. Model "DEX.VF3.E7" with 13 parameters had minimum AIC, though the corresponding log $\mathcal{L}$ was not very different from that for "LP.R4r2.E7". Involving a more stringent penalty for the number of parameters fitted, BIC suggested that "EXP.VF3.E1" with only six parameters sufficed to model the covariance structure adequately.

Estimated variances for selected analyses are shown in Figure 1, and corresponding correlations are given in Figure 2. For $e = 1$ and $v \geq 4$ or

---

[1] available as part of DFREML 3.0 from http://agbu.une.edu.au/~kmeyer/dfreml.html
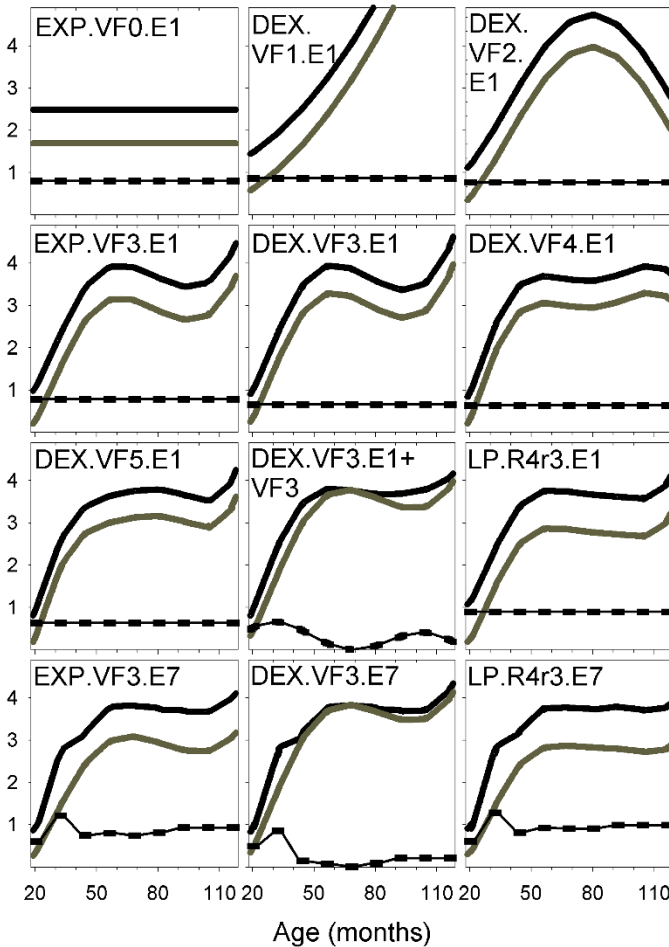
Figure 1. Estimates of variances (Solid black line: phenotypic, grey line: animal and dotted line: measurement error, with "thick" areas of line representing ages in the data; in 1 000 kg^2) from phenotypic analyses.

$\nu = 3$ together with heterogeneous σ_ε^2, estimates of the total variance differed little between models of analyses. Differences were largest at the highest ages which were represented by comparatively few records only. For DEX, however, estimates of σ_ε^2 were consistently lower than for the other models, in particular if a VF $\mathcal{E}$ was fitted or $e > 1$.

As shown in Figure 2, correlation functions EXP and DEX force estimates of within-animal correlations between repeated records to decrease steadily with increasing lag. Corresponding phenotypic correlation (r_P) estimates follow a similar pattern, but fluctuate if variances are heterogeneous.

Table II. Estimates of variance components at mean age [a] and parameters of correlation functions [b] for **phenotypic** analyses, together with maximum likelihood values ($\log \mathcal{L}$; $+13\,300$), and Akaike's (AIC) and Schwarz' Bayesian (BIC) Information Criterion (both $-26\,600$).

<div align="right">(continued on the next page)</div>

Model [c]	p [d]	$\log \mathcal{L}$	AIC	BIC	σ^2_{R0}	$\sigma^2_{\varepsilon 0}$	σ_{P0}	ρ	θ	λ
					Homogeneous measurement error variance					
CS.VF0.E1	3	−307.2	620.4	638.5	1505	1024	2529	0.8918		
EXP.VF0.E1	3	−263.4	532.8	550.9	1688	804	2492	0.9894	0.0107	1.0000
GAU.VF0.E1	3	−271.2	548.3	566.5	1485	1002	2488	0.9998	0.0002	2.0000
DEX.VF0.E1	4	−261.9	531.7	556.0	2488	6	2494	0.8192	0.1995	0.3777
DEX.VF1.E1	5	−165.0	340.0	370.3	3208	862	4069	0.9969	0.0031	1.0755
CS.VF2.E1	5	−136.6	283.2	313.5	3738	764	4503	0.8755		
EXP.VF2.E1	5	−113.8	237.5	267.8	3712	817	4529	0.9956	0.0044	1.0000
DEX.VF2.E1	6	−112.8	237.5	273.9	3775	764	4539	0.9871	0.0129	0.7334
SAD.VF2.E1	6	−112.9	237.9	274.2	3668	839	4507	0.9999	0.0001	1.8715
CS.VF3.E1	6	−115.8	243.6	280.0	3284	638	3922	0.8235		
EXP.VF3.E1	6	−95.4	202.8	239.1	3114	778	3892	0.9946	0.0054	1.0000
GAU.VF3.E1	6	−108.3	228.6	264.9	2977	922	3899	0.9999	0.0001	2.0000
DEX.VF3.E1	7	−91.7	197.4	239.9	3201	656	3857	0.9689	0.0316	0.5681
SAD.VF3.E1	7	−94.8	203.6	246.0	3089	794	3883	0.9995	0.0005	1.5448
EXP.VF4.E1	7	−92.6	199.1	241.5	2918	769	3687	0.9944	0.0056	1.0000
DEX.VF4.E1	8	−88.2	192.4	240.9	2981	633	3614	0.9642	0.0365	0.5415
DEX.VF5.E1	9	−86.9	**191.8**	246.4	3124	629	3753	0.9624	0.0384	0.5275
DEX.VF6.E1	10	−86.5	193.1	253.7	3116	630	3745	0.9636	0.0371	0.5394
DEX.VF7.E1	11	−86.5	195.0	261.7	3121	629	3751	0.9637	0.0370	0.5400
LP.R3r1.E1	4	−145.6	299.2	323.4	3292	1032	4324			
LP.R3r2.E1	6	−119.6	251.2	287.6	3232	952	4184			
LP.R4r2.E1	8	−99.8	215.6	264.0	2824	917	3741			
LP.R4r3.E1	10	−95.7	211.4	272.0	2842	891	3733			
LP.R5r3.E1	13	−90.0	206.0	284.8	2871	854	3725			

Table II. *continued.*

Model[c]	p[d]	log$\mathcal{L}$	AIC	BIC	σ^2_{R0}	$\sigma^2_{\varepsilon 0}$	σ_{P0}	ρ	θ	λ
				Measurement error variance function						
DEX.VF2.E1+VF3	9	−91.4	200.7	255.3	4171	10	4181	0.9092	0.0952	0.3601
DEX.VF3.E1+VF3	10	−82.1	184.2	**244.8**	3760	1	3761	0.8967	0.1091	0.3341
DEX.VF4.E1+VF3	11	−81.1	184.2	250.8	3648	5	3653	0.8851	0.1221	0.3246
DEX.VF4.E1+VF4	12	−78.9	**181.8**	254.5	3782	20	3802	0.8942	0.1118	0.3320
LP.R3.E1+VF3	10	−102.0	223.9	284.5	3148	979	4127			
LP.R3.E1+VF4	11	−95.7	213.3	280.0	3128	864	3992			
LP.R3.E1+VF5	12	−93.7	211.5	284.2	3133	906	4039			
LP.R4r3.E1+VF3	13	−82.9	191.9	270.6	2829	912	3742			
LP.R4r3.E1+VF4	14	−79.6	187.2	272.1	2857	822	3679			
				Annual (seven) measurement error variances						
EXP.VF3.E7	12	−77.3	178.7	**251.4**	3073	739	3812	0.9958	0.0042	1.0000
DEX.VF3.E7	13	−74.8	**175.6**	254.4	3819	0	3819	0.8879	0.1189	0.3224
EXP.VF4.E7	13	−77.1	180.3	259.1	3035	744	3780	0.9958	0.0043	1.0000
DEX.VF4.E7	14	−74.7	177.4	262.2	3780	0	3780	0.8857	0.1214	0.3199
DEX.VF5.E7	15	−74.7	179.4	270.3	3783	0	3783	0.8863	0.1207	0.3205
LP.R3r2.E7	12	−88.7	201.3	274.0	3147	922	4069			
LP.R4r2.E7	14	−74.5	177.0	261.8	2857	916	3773			
LP.R5r2.E7	16	−74.4	180.9	277.8	2878	915	3793			

(a) σ^2_{R0}: estimate of permanent environmental variance at mean; $\sigma^2_{\varepsilon 0}$: estimate of measurement error variance.

(b) ρ: correlation parameter for CS and SAD, corresponding value calculated from θ for EXP, GAU and DEX; θ: exponential parameter, and λ: scaling parameter.

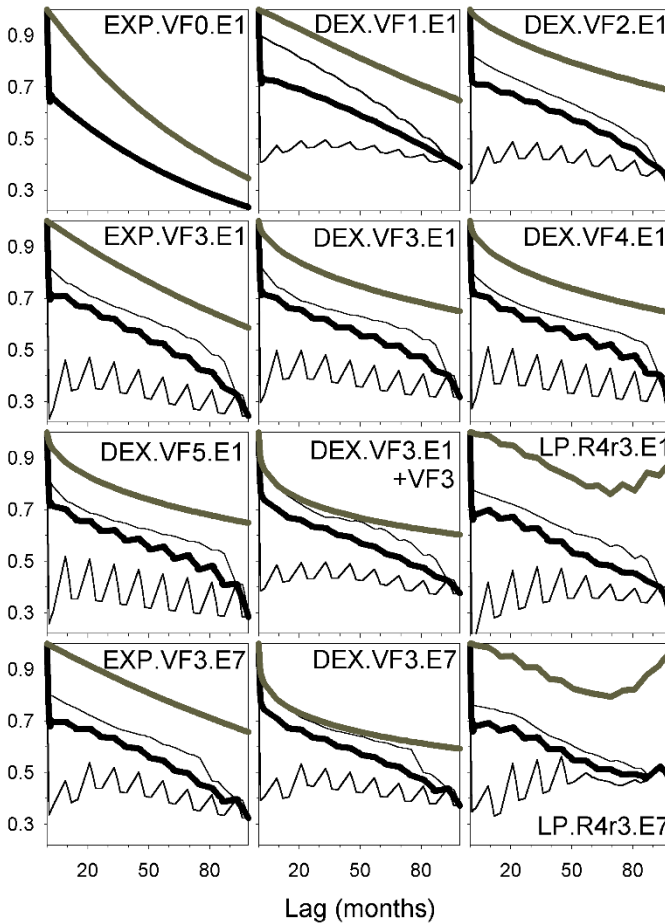(c) See text.

(d) No. of parameters.

Figure 2. Estimates of (average) correlations for lags in age from phenotypic analyses (Grey line: animal, thick black line: phenotypic, thin black lines showing ranges for phenotypic correlations).

Hence average values together with their ranges are shown for r_P. Whilst EXP and DEX imply stationarity for r_R, analyses fitting LPs do not impose this restriction, and values for r_R shown are averages (with ranges), as for r_P. Previous analyses fitting LPs [23] had reported an inexplicable peak in estimates or r_R between the youngest and oldest ages, and attributed this to sampling variation. This peak is reflected in increases of the average estimate of r_R for lags of more than 70 months for analyses "LP.R4r3.E1" and "LP.R4r3.E7". Similarity in values of $\log \mathcal{L}$ and estimates of variances for the latter and "DEX.VF3.E7" suggests that differences in estimates of r_R are indeed spurious.
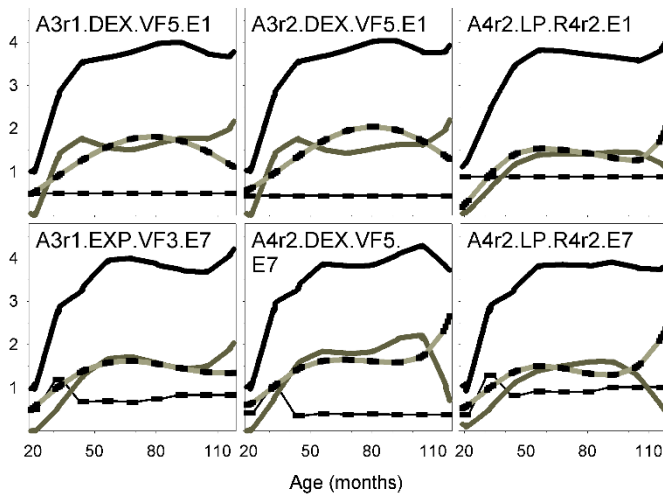
Figure 3. Estimates of variances components from *genetic* analyses (Thick black line: phenotypic, black and grey line: genetic, grey line: permanent environmental and dotted line: measurement error, with "thick" areas representing ages in the data ; in $1\,000\ \text{kg}^2$).

4.3.2. Genetic analyses

Table III gives results from analyses fitting a set of RR coefficients for animals' additive genetic effects in addition. Clearly, there is variation between individuals which should not be ignored. Likelihoods increased up to $k_A = 4$, the order of fit identified previously [23]. Fitting $\mathbf{K}_A$ with rank $k'_A = 2$ was sufficient throughout. Again, BIC favoured a model with homogeneous σ^2_ε, "A3r1.DEX.VF5.E1" with 12 parameters, while AIC selected the model of previous analyses, "A4r2.LP.R4r2.E7" with 21 parameters.

Estimates of variances and correlations for the "best" models and others closely related to them are shown in Figures 3 and 4, respectively. For comparison, corresponding estimates from a "standard" multivariate analysis treating observations at each year of age (2 to 10) as a different trait are given in Figure 5. Most notable again is the similarity in estimates of the total variance for all analyses shown while the partitioning into invidual sources of variation differs. Problems with sharply diverging estimates of genetic and permanent environmental variances for the highest ages in this data set were observed before and attributed to the influence of atypical records at the upper extreme.

Similarly, estimates of the average r_P (Fig. 4) for the analyses shown agree well, ranging from 0.7 to 0.8 to about 0.4. Estimating $\mathbf{K}_A$ with rank $k'_A = 1$ forced all estimates of the genetic correlation (r_A) between records at different ages to be unity, while $k'_A = 2$ allowed for a decline in r_A with increasing lag.

Table III. Estimates of correlation function parameters [a] and variance components [b] at mean age, together with maximum likelihood values (log $\mathcal{L}$; +13 300) and Akaike's (AIC) and Schwarz' Bayesian (BIC) Information Criterion (both −26 600), for **genetic** analyses fitting different correlation structures.

(continued on the next page)

Model [c]		p [d]	$\log\mathcal{L}$	AIC	BIC	σ^2_{A0}	σ^2_{R0}	$\sigma^2_{\varepsilon 0}$	σ^2_{P0}	ρ	θ	κ
					Homogeneous measurement error variances							
A1	DEX.VF3.E1	8	−70.2	156.4	204.9	448	2651	554	3653	0.9487	0.0526	0.5583
A1	DEX.VF5.E1	10	−56.8	133.5	194.1	642	2235	411	3288	0.9179	0.0857	0.5119
A1	DEX.VF7.E1	12	−55.4	134.7	207.4	648	2351	399	3398	0.9142	0.0897	0.5032
A1	SAD.VF7.E1	12	−65.2	154.5	227.2	582	2202	531	3314	0.8150	0.2045	0.3947
A2	DEX.VF5.E1	12	−49.5	123.1	195.8	1296	1772	414	3482	0.9003	0.1050	0.5060
A2	DEX.VF7.E1	14	−47.8	123.6	208.4	1297	1885	411	3593	0.8948	0.1112	0.4863
A3r1	DEX.VF3.E1	10	−59.7	139.5	200.1	1595	1769	611	3975	0.9510	0.0503	0.6680
A3r1	DEX.VF4.E1	11	−57.9	137.8	204.4	2131	1175	570	3875	0.9467	0.0548	0.7537
A3r1	EXP.VF5.E1	11	−50.8	123.6	190.2	2054	1239	581	3875	0.9739	0.0265	1.0000
A3r1	DEX.VF5.E1	12	−46.2	116.3	**189.0**	1764	1512	499	3776	0.9229	0.0803	0.6044
A3r2	DEX.VF3.E1	12	−58.3	140.6	213.3	1720	1676	575	3970	0.9288	0.0738	0.5714
A3r2	EXP.VF5.E1	13	−48.8	123.7	202.4	2389	1034	517	3940	0.9641	0.0366	1.0000
A3r2	DEX.VF5.E1	14	−42.9	**113.9**	198.7	1955	1434	440	3829	0.8908	0.1157	0.5278
A3r2	SAD.VF5.E1	14	−48.4	124.9	209.7	2432	975	511	3919	0.9386	0.0634	0.8713
A3r2	EXP.VF6.E1	14	−48.7	125.4	210.2	2421	988	516	3925	0.9632	0.0375	1.0000
A3r2	EXP.VF7.E1	15	−48.7	127.5	218.3	2402	1009	516	3927	0.9639	0.0368	1.0000
A3r2	LP.R4r2.E1	13	−71.8	169.5	248.3	1525	1433	882	3841			
A4r1	DEX.VF3.E1	11	−58.8	139.5	206.2	1600	1718	635	3954	0.9561	0.0449	0.7116
A4r1	EXP.VF5.E1	12	−49.1	122.1	194.8	2060	1133	588	3782	0.9718	0.0286	1.0000
A4r1	DEX.VF5.E1	13	−45.8	117.6	196.3	1792	1423	516	3732	0.9277	0.0750	0.6354
A4r1	DEX.VF7.E1	15	−44.9	119.9	210.8	1848	1408	525	3782	0.9259	0.0770	0.6348
A4r2	DEX.VF3.E1	14	−55.6	139.2	224.1	1695	1579	614	3888	0.9220	0.0812	0.5101
A4r2	EXP.VF5.E1	15	−46.9	123.8	214.7	2279	1067	521	3868	0.9629	0.0378	1.0000
A4r2	DEX.VF5.E1	16	−42.1	116.1	213.0	1995	1350	433	3779	0.8951	0.1108	0.5725
A4r2	DEX.VF7.E1	18	−40.5	117.0	226.1	1954	1448	420	3822	0.8904	0.1161	0.5575
A4r2	LP.R4r2.E1	15	−65.7	161.4	252.3	1500	1407	883	3790			

Table III. *continued.*

Model [c]		p [d]	$\log \mathcal{L}$	AIC	BIC	σ_{A0}^2	σ_{R0}^2	$\sigma_{\varepsilon 0}^2$	σ_{P0}^2	ρ	θ	κ
					Heterogeneous measurement error variances							
A3r1	EXP.VF3.E7	15	−40.2	110.3	**201.2**	1585	1750	674	4009	0.9889	0.0112	1.0000
A3r1	DEX.VF3.E7	16	−39.5	111.1	208.0	1539	2043	439	4021	0.9502	0.0511	0.6400
A3r1	DEX.VF5.E7	18	−37.4	110.8	219.9	1646	1891	332	3869	0.9097	0.0946	0.5335
A3r1	LP.R4r2.E7	17	−39.6	113.2	216.2	1323	1673	901	3897			
A3r2	EXP.VF3.E7	17	−39.6	113.2	216.1	1674	1660	675	4008	0.9886	0.0115	1.0000
A3r2	DEX.VF3.E7	18	−37.7	111.5	220.5	1787	1653	562	4001	0.9655	0.0351	0.7289
A3r2	EXP.VF5.E7	19	−36.6	111.3	226.4	1804	1401	701	3905	0.9881	0.0120	1.0000
A3r2	DEX.VF5.E7	20	−35.7	111.4	232.6	1797	1666	412	3874	0.9243	0.0787	0.5811
A3r2	LP.R4r2.E7	19	−34.4	106.9	222.0	1562	1447	890	3899			
A4r1	EXP.VF3.E7	16	−39.6	111.3	208.2	1617	1704	661	3982	0.9879	0.0122	1.0000
A4r1	DEX.VF3.E7	17	−38.6	111.3	214.3	1557	2413	10	3980	0.8550	0.1566	0.4305
A4r1	DEX.VF5.E7	19	−36.6	111.3	226.4	1682	1927	245	3854	0.8843	0.1229	0.4975
A4r2	EXP.VF3.E7	19	−36.0	110.1	225.2	1543	1712	684	3939	0.9893	0.0107	1.0000
A4r2	DEX.VF3.E7	20	−35.0	110.1	231.2	1548	2314	61	3923	0.8497	0.1628	0.4030
A4r2	SAD.VF3.E7	20	−36.0	112.1	233.3	1567	1691	678	3935	0.9890	0.0111	1.0045
A4r2	DEX.VF5.E7	22	−33.3	110.6	243.9	1645	1794	378	3817	0.9122	0.0919	0.5388
A4r2	LP.R4r2.E7	21	−31.2	**104.5**	231.7	1443	1509	895	3846			
A4r3	DEX.VF5.E7	24	−32.4	112.9	258.3	1789	1745	293	3828	0.8761	0.1323	0.4654
A4r3	LP.R4r2.E7	23	−30.6	107.3	246.6	1506	1439	891	3836			

(a) ρ: correlation parameter for SAD (corresponding value calculated from θ for EXP and DEX), θ: exponential parameter for EXP and DEX (corresponding value calculated from ρ for SAD), and κ: scaling parameter for DEX and SAD.

(b) σ_{A0}^2: additive genetic variance, σ_{R0}^2: permanent environmental variance, $\sigma_{\varepsilon 0}^2$: measurement error variance, and σ_{P}^2: phenotypic variance.
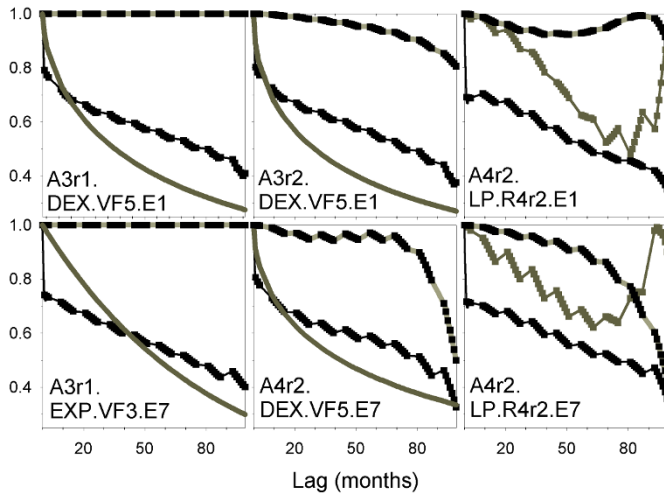
(c) See text.

(d) No. of parameters.

Figure 4. Estimates of (average) correlations for lags in age from *genetic* analyses (Grey line: permanent environmental, black and grey line: genetic, and black line: phenotypic).
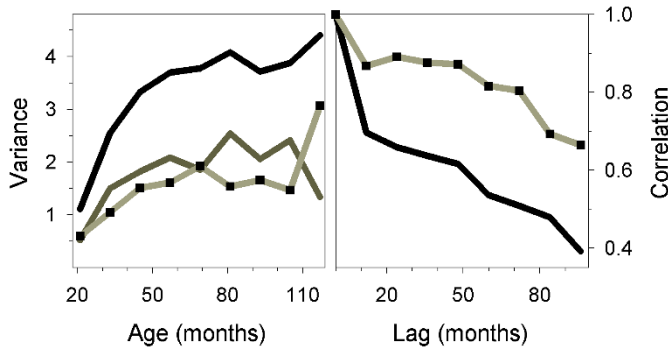

Figure 5. Estimates of variances (left; in $1\,000$ kg^2) and average correlations (right) from *multivariate* analyses treating observations for each year of age as separate traits (black line: phenotypic, black and grey line: genetic, and grey line: residual).

5. DISCUSSION

Whilst analyses of longitudinal, spatial or similar data assuming a parametric correlation structure or covariance function are commonplace in other areas of applied statistics, they have found few applications in the analysis of animal breeding data, *e.g.* [1, 37]. Harville [10] considered auto-regressive random effects in linear mixed models, and suggested that modelling of dairy records might be improved by considering permanent environmental effects of cows

as auto-regressive rather than constant across lactations. "Random coefficient" models are often found to be of little advantage and somewhat cumbersome, the resulting covariance functions generally yielding less readily interpretable results than the parameters of a parametric correlation function. Moreover, low degree polynomial regressions tend to provide a poor fit for trajectories with steep initial increases which then level off to gradually approach an asymptote [3], p. 102, *i.e.* high order polynomials and corresponding numbers of parameters are often required to model the covariance structure for growth curve type of analyses.

In contrast, animal breeders have embraced random regression models for the analysis of longitudinal data, in particular test-day records of dairy cows and growth data for pigs and beef cattle. This can be attributed to several factors.

Firstly, quantitative genetic analyses are invariably concerned with the variation between animals, while other areas of statistics are often content with modelling within-subject covariances only. Fitting a set of additive genetic RR coefficients provides estimates of genetic merit for the whole range of ages considered, and allows ranking of animals to change with time. RR models are thus an obvious choice if we are concerned with (genetic) differences between individuals. Assuming a RR model and resulting covariance structure on a genetic level, it seems natural to apply the same model to other random effects, for instance permanent environmental effects due to the individual.

Secondly, parametric models are usually formulated assuming homogeneous variances. Variances of "repeated" record traits of interest to animal breeders, however, often change with time, if only due to scale effects. While extensions to heterogeneous models by replacing single variances with a (polynomial) VF are straightforward, these are seldom described. RR models account for changes in variances with time, and for RR involving orthogonal polynomials, do not require any specific assumptions about the shape of the resulting VF.

Thirdly, estimates of genetic covariance matrices arising from RR model analyses can be thought of as smoothed versions of corresponding estimates from an unstructured, multivariate analysis treating records at different ages as different traits. Covariance functions which give covariances between records at individual ages as function of orthogonal polynomials of the ages and the elements of a matrix of coefficients ($\mathbf{K}$), have been described as "infinite-dimensional" equivalent to covariance matrices in standard multivariate analyses. Estimates of the eigenvalues and eigenfunctions of such CFs can be obtained directly as the eigenvalues of $\mathbf{K}$ and from the corresponding eigenvectors. For genetic covariance functions, these statistics provide valuable insight into the effects of selection for the trait considered. See Kirkpatrick *et al.* [16, 17] for further details.

Last, but not least, RR models provide a computationally feasible way to estimate CF for large data sets with records coming in at "all ages", as are typical for data from livestock recording schemes. Estimating **K** as the matrix of covariances between RR coefficients requires mixed model equations of size proportional to the number of regression coefficients to be manipulated, rather than proportional to the number of ages or even the number of records. In contrast, general algorithms for the analysis of longitudinal data fitting parametric CF or RF are frequently formulated in terms of the variance matrix of the vector of observations, *i.e.* involve the latter, *e.g.* [3], Chap. 4.

This paper shows how the preferred RR approach for genetic effects can be combined with a parametric correlation model for within animal covariances. It extends the model suggested by Foulley *et al.* [5] to a variety of correlation structures, both stationary and non-stationary, and heterogeneous variances. An average information algorithm for the estimation of the parameters for the covariance, variance and correlation functions describing the dispersion structure for such model, by REML is outlined. It involves computations proportional to the number of animals and the order of polynomial fit for genetic effects. Extensions to models involving additional sets of RR coefficients for other random effects, *e.g.* maternal effects, or more complicated correlation functions are straightforward.

Application to a data set of mature weight of beef cows showed that assuming a parametric correlation structure for within animal covariances can result in a considerably more parsimonious model than a RR model for permanent environmental effects. The example showed further that parametric correlation functions can eliminate erratic and inexplicable estimates of correlations between records at extreme ages which have been encountered with a RR. Some differential partitioning of the total variance was evident though. This occurred in particular, when both temporary and permanent environmental variances were assumed heterogeneous. An alternative for this scenario might be the use of a link function as suggested by Foulley *et al.* [6], which would assume a functional relationship between the two sources of variation and thus might not only reduce the number of parameters required further but also alleviate problems of erratic partitioning. Further work is required to examine the effect of different assumptions about the structure of within-animal covariances on the partitioning of the phenotypic variance and, thus, their potential impact on estimates of genetic parameters and predictions of breeding values.

REFERENCES

[1] Carvalheira J.G.V., Blake R.W., Pollak E.J., Quaas R.L., Duran-Castro C.V., Application of an autoregressive process to estimate genetic parameters and breeding values for daily milk yield in a tropical herd of Lucerna cattle and in United States Holstein herds, J. Dairy Sci. 81 (1998) 2738–2751.

[2] Diggle P.J., An approach to the analysis of repeated measurements, Biometrics 44 (1988) 959–971.

[3] Diggle P.J., Liang K.Y., Zeger S.L., Analysis of Longitudinal Data, Oxford Science Publications, Clarendon Press, Oxford, 1994.

[4] Eubank R.L., Speckman P., Curve fitting by polynomial-trigonometric regression analysis, Biometrika 77 (1990) 1–9.

[5] Foulley J.-L., Jaffrézic F., Robert-Granié C., EM-REML estimation of covariance parameters in Gaussian mixed models for longitudinal data analysis, Genet. Select. Evol. 32 (2000) 129–141.

[6] Foulley J.-L., Quaas R.L., Thaon d'Arnoldi C., A link function approach to heterogeneous variance components, Genet. Select. Evol. 30 (1998) 27–44.

[7] Fuller W.A., Grafted polynomials as approximating functions, Austr. J. Agric. Econ. 13 (1969) 35–46.

[8] Gilmour A.R., Thompson R., Cullis B.R., Average Information REML, an efficient algorithm for variance parameter estimation in linear mixed models, Biometrics 51 (1995) 1440–1450.

[9] Graser H.-U., Smith S.P., Tier B., A derivative-free approach for estimating variance components in animal models by restricted maximum likelihood, J. Anim. Sci. 64 (1987) 1362–1370.

[10] Harville D.A., Recursive estimation using mixed linear models with autoregressive random effects, in: Van Vleck L.D., Searle S.R. (Eds.), Variance Components and Animal Breeding, Cornell University, Proceedings Conference in Honor of Henderson C.R., Ithaca, NY, 1979, pp. 157–179.

[11] Huggins R.M., Loesch D.Z., On the analysis of mixed longitudinal growth data, Biometrics 54 (1998) 583–595.

[12] Jaffrézic F., White I.M.S., Thompson R., Hill W.G., A link function approach to model heterogeneity of residual variances over time in lactation curve analyses, J. Dairy Sci. 83 (2000) 1089–1093.

[13] Jennrich R.I., Schluchter M.D., Unbalanced repeated-measures models with structured covariance matrices, Biometrics 42 (1986) 805–820.

[14] Johnson D.L., Thompson R., Restricted maximum likelihood estimation of variance components for univariate animal models using sparse matrix techniques and average information, J. Dairy Sci. 78 (1995) 449–456.

[15] Kenward M.G., A method for comparing profiles of repeated measurements, Appl. Stat. 36 (1987) 296–308.

[16] Kirkpatrick M., Hill W.G., Thompson R., Estimating the covariance structure of traits during growth and aging, illustrated with lactations in dairy cattle, Genet. Res. 64 (1994) 57–69.

[17] Kirkpatrick M., Lofsvold D., Bulmer M., Analysis of the inheritance, selection and evolution of growth trajectories, Genetics 124 (1990) 979–993.

[18] Lindstrom M.J., Bates D.M., Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data, J. Amer. Stat. Ass. 83 (1988) 1014–1022.

[19] Madsen P., Jensen J., Thompson R., Estimation of (co)variance components by REML in multivariate mixed linear models using average of observed and expected information, in: Proceedings Fifth World Congress on Genetics Applied to Livestock Production, 7–12 August 1994, University of Guelph, Ontario, Vol. 22, 1994, pp. 19–22.

[20] Meyer K., An "average information" restricted maximum likelihood algorithm for estimating reduced rank genetic covariance matrices or covariance functions for animal models with equal design matrices, Genet. Select. Evol. 29 (1997) 97–116.

[21] Meyer K., "DxMrr" – a program to estimate covariance functions for longitudinal data by restricted maximum likelihood., in: Proceedings Sixth World Congress on Genetics Applied to Livestock Production, 12–16 January 1998, University of New England, Armidale, Vol. 27, 1998, pp. 465–466.

[22] Meyer K., Estimating covariance functions for longitudinal data using a random regression model, Genet. Select. Evol. 30 (1998) 221–240.

[23] Meyer K., Estimates of genetic and phenotypic covariance functions for postweaning growth and mature weight of beef cows, J. Anim. Breed. Genet. 116 (1999) 181–205.

[24] Meyer K., Random regressions to model phenotypic variation in monthly weights of Australian beef cows, Livest. Prod. Sci. 65 (2000) 19–38.

[25] Meyer K., Estimates of direct and maternal covariance functions for growth of Australian beef calves from birth to weaning, Genet. Select. Evol. 33 (2001) 487–514.

[26] Meyer K., Carrick M.J., Donnelly B.J.P., Genetic parameters for growth traits of Australian beef cattle from a multi-breed selection experiment, J. Anim. Sci. 71 (1993) 2614–2622.

[27] Meyer K., Hill W.G., Estimation of genetic and phenotypic covariance functions for longitudinal data by restricted maximum likelihood, Livest. Prod. Sci. 47 (1997) 185–200.

[28] Meyer K., Smith S.P., Restricted maximum likelihood estimation for animal models using derivatives of the likelihood, Genet. Select. Evol. 28 (1996) 23–49.

[29] Muñoz A., Carey V., Schouten J.P., Segal M., Rosner B., A parametric family of correlation structures for the analysis of longitudinal data, Biometrics 48 (1992) 733–742.

[30] Nelder J.A., Mead R., A simplex method for function minimization, Computer J. 7 (1965) 308–313.

[31] Núñez-Antón V., Woodworth G.G., Analysis of longitudinal data with unequally spaced observations and time-dependent correlated errors, Biometrics 50 (1994) 445–456.

[32] Núñez-Antón V., Zimmerman D.L., Modelling non-stationary longitudinal data, Biometrics 56 (2000) 699–705.

[33] Pletcher S.D., Geyer C.J., The genetic analysis of age-dependent traits: Modeling the character process, Genetics 153 (1999) 825–835.

[34] Pourahmadi M., Joint mean-covariance models with applications to longitudinal data: Unconstrained parameterisation, Biometrika 86 (1999) 677–690.

[35] Schnyder U., Hofer A., Labroue L., Kuenzi N., Genetic parameters of a random regression model for daily feed intake of performance tested French Landrace and Large White growing pigs, Genet. Select. Evol. 33 (2001) 635–658.

[36] Smith S.P, Differentiation of the Cholesky algorithm, J. Comp. Graph. Stat. 4 (1995) 134–147.

[37] Wade K.M., Quaas R.L., Van Vleck L.D., Estimation of the parameters involved in a first order autoregressive process for contemporary groups, J. Dairy Sci. 76 (1993) 3033–3040.

[38] Wolfinger R.D., Covariance structure selection in general mixed models, Comm. Stat. - Simul. Comp. 22 (1993) 1079–1106.

[39] Zimmerman D.L., Núñez-Antón V., ElBarmi H., Computational aspects of likelihood based estimation of first-order antedependence models, J. Stat. Comp. Simul. 60 (1998) 67–84.

[40] Zimmermann D.L., Harville D.A., A random field approach to the analysis of field-plot experiments and other spatial experiments, Biometrics 47 (1991) 223–239.