

ESTIMATES OF GENETIC COVARIANCE FUNCTIONS FOR WEIGHTS OF BEEF COWS ASSUMING A PARAMETRIC CORRELATION STRUCTURE FOR ENVIRONMENTAL EFFECTS

Karin Meyer

Animal Genetics and Breeding Unit¹, University of New England, Armidale, NSW 2351

SUMMARY

Repeated records for mature weights of beef cows were analysed fitting a random regression model for genetic effects whilst accounting for permanent environmental covariances between records for an animal through correlated residuals. Changes in these covariances were modelled by a polynomial variance function combined with a parametric correlation function. Results show such models to be well suited to describe the covariance structure among repeated records. They provide a more parsimonious alternative to models fitting random regression coefficients for genetic and permanent environmental effects, and reduce problems of erratic correlation estimates for extreme ages.

Keywords: Mature weight, beef cows, correlation function, parametric

INTRODUCTION

Random regression (RR) analyses fitting a polynomial regression on time have become popular for the analysis of longitudinal data in animal breeding applications. Typically, RR analyses model both genetic and permanent environmental (p.e.) effects due to animals through such functions, which yield covariance functions (CF) determined by the covariances among the RR coefficients. In contrast, other areas of applied statistics generally model 'within-subject' covariances among repeated records through a correlation function (RF) and corresponding variance(s). This paper presents an analysis of mature weight records of beef cow, combining a regression on polynomials of age to model changes in additive genetic effects and covariances over time with a parametric RF and variance function (VF) to model p.e. covariances between records for the same animal.

MATERIAL AND METHODS

Data. Data consisted of January weights of Polled Hereford cows in the Wokalup selection experiment, which tended to represent cows at the their top weight during the year. There were 3320 records on 850 cows, with 50, 153, 103, 124, 83, 75, 47, 50 and 50 animals having 1, 2, ..., 8 and 9 records, respectively.

With the bulk of calves born in April and May each year, ages at weighing - ranging from 19 to 119

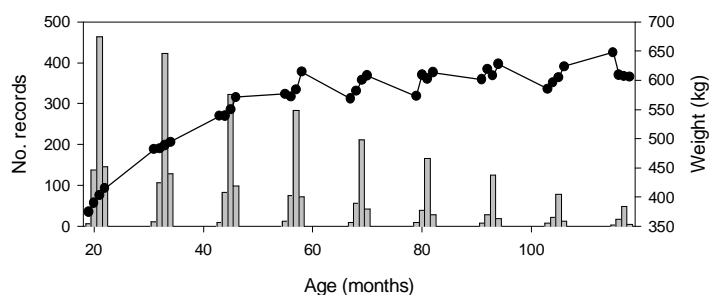


Figure 1. Numbers of records (bars) and mean weights (●).

¹ *AGBU is a joint Institute of NSW Agriculture and the University of New England.*

months - were not distributed continuously. Figure 1 gives numbers of records and mean weights for the ages in the data. Records have been analysed previously fitting RR on Legendre polynomials of age for both genetic and p.e. effects (Meyer 1999).

Analysis. Restricted maximum likelihood analyses were carried out fitting a RR on Legendre polynomials (LP) of age at recording (in months) for animals' genetic effects (A), estimating the corresponding matrix of covariances among RR coefficients. Orders of fit up to $k=4$ were considered. P.e. covariances between repeated records for an animal were modelled through a RF with 1 or 2 parameters and a polynomial VF with orders of fit up to $v=7$. Correlations were assumed to follow compound symmetry (CS), an exponential or power function (EXP), a 'damped' exponential (DEX), and a first-order structured ante-dependence (SAD) model (Nuñez-Antón and Zimmerman 2000). Expressions for r_{ij} - the correlation between records taken at times t_i and t_j with $t_i < t_j$ and $r_{ji} = r_{ij}$ - are given in Table 1. Whilst the first three RF are stationary, i.e. depend only on the lag between observations, the SAD facilitates non-stationarity through a deformation of the time scale, $f(t_i, \kappa) = (t_i^\kappa - 1)/\kappa$ for $\kappa \neq 0$ and $\log(t_i)$ otherwise. In addition, p.e. covariances were modelled through RR on LP of age. As in previous analyses, measurement error variances were considered to be either

Table 1. Correlation functions.

RF	P ^A	r_{ij}
CS	ρ	1 for $i=j$, ρ for $i \neq j$
EXP	θ	$e^{-\theta(t_j - t_i)}$
DEX	θ, κ	$e^{-\theta(t_j - t_i)^\kappa}$
SAD	ρ, κ	$\rho^{f(t_j, \kappa) - f(t_i, \kappa)}$ for $j=i+1$ $\prod_{k=i+1}^{j-1} \rho^{f(t_k, \kappa) - f(t_{k+1}, \kappa)}$ for $j > i+1$

^A Parameters – see results

homogeneous ($e=1$) or different for each year of age (treating years 8-10 as single class), resulting in $e=7$ variances. Estimated covariances among RR were forced to have reduced rank if eigenvalues were close to zero. In the following, 'Akrm' and 'LP.Rkrm' denote analyses fitting Legendre polynomial based CF of order k and rank m for genetic and p.e. effects, respectively, and 'XXX.VFv' is a model fitting a parametric RF together with VF of order v . Models were compared using penalised likelihoods, i.e. the REML forms of Akaike's (AIC) and Schwarz' Bayesian (BIC) information criteria.

RESULTS

Table 1 summarises likelihood criteria (scaled) and estimated parameters (ρ : auto-correlation, θ : corresponding exponential parameter, and κ : scale parameter) of the RF for selected analyses. Allowing for covariance matrices among RR coefficients of rank greater than 2 (not shown) did not yield any significant increases log likelihood (L). Assuming homogeneous error variances, an increase in order of polynomial fit to model p.e. variances increased log L markedly, up to $v=5$, whilst $v=3$ sufficed when allowing for heterogeneous error variances. Similarly, the damped exponential RF, DEX, was advantageous over EXP for $e=1$ but less so for $e=7$, showing how the partitioning of the total environmental variance was highly dependent on the assumptions made. Fitting a SAD gave similar or lower likelihood values than for a stationary RF with the same number of parameters, i.e. there was no evidence for non-stationarity in the data. Estimates of $\rho = \exp\{-\theta\}$ were consistently high. For $e=1$, log L was higher and information criteria were substantially lower for models fitting a parametric RF together with a VF for p.e. variances than for analyses fitting LP based CF. Presumably, this was due to a greater scope for the VF to account for changes in

Table 2. Likelihoods^A and estimates of correlation function.

Model		p ^B	log L	AIC	BIC	ρ	θ	κ
Homogeneous error variance ($e=1$)								
A1	DEX.VF3	8	-40.2	16.4	24.9	0.949	0.0526	0.558
A1	DEX.VF5	10	-26.8	33.5	14.1	0.918	0.0857	0.512
A2	DEX.VF5	12	-19.5	23.1	15.8	0.900	0.1050	0.506
A3r1	DEX.VF4	11	-27.9	37.8	24.4	0.947	0.0548	0.754
A3r1	EXP.VF5	11	-20.8	23.6	10.2	0.974	0.0265	1.000
A3r1	DEX.VF5	12	-16.2	16.3	9.0	0.923	0.0803	0.604
A3r2	EXP.VF5	13	-18.8	23.7	22.4	0.964	0.0366	1.000
A3r2	DEX.VF5	14	-12.9	13.9	18.7	0.891	0.1157	0.528
A3r2	SAD.VF5	14	-18.4	24.9	29.7	0.939	0.0634	0.871
A3r2	LP.R4r2	13	-41.8	69.5	68.3			
A4r1	EXP.VF5	12	-19.1	22.1	14.8	0.972	0.0286	1.000
A4r1	DEX.VF5	13	-15.8	17.6	16.3	0.928	0.0750	0.635
A4r2	EXP.VF5	15	-16.9	23.8	34.7	0.963	0.0378	1.000
A4r2	DEX.VF5	16	-12.1	16.1	33.0	0.895	0.1108	0.573
A4r2	LP.R4r2	15	-35.7	61.4	72.3			
Heterogeneous error variances ($e=7$)								
A3r1	EXP.VF3	15	-10.2	10.3	21.2	0.989	0.0112	1.000
A3r1	DEX.VF3	16	-9.5	11.1	28.0	0.950	0.0511	0.640
A3r1	DEX.VF5	18	-7.4	10.8	39.9	0.910	0.0946	0.534
A3r1	LP.R4r2	17	-9.6	13.2	36.2			
A3r2	EXP.VF3	17	-9.6	13.2	36.1	0.989	0.0115	1.000
A3r2	DEX.VF3	18	-7.7	11.5	40.5	0.965	0.0351	0.729
A3r2	EXP.VF5	19	-6.6	11.3	46.4	0.988	0.0120	1.000
A3r2	DEX.VF5	20	-5.7	11.4	52.6	0.924	0.0787	0.581
A3r2	LP.R4r2	19	-4.4	6.9	42.0			
A4r1	EXP.VF3	16	-9.6	11.3	28.2	0.988	0.0122	1.000
A4r1	DEX.VF3	17	-8.6	11.3	34.3	0.855	0.1566	0.431
A4r1	DEX.VF5	19	-6.6	11.3	46.4	0.884	0.1229	0.498
A4r2	EXP.VF3	19	-6.0	10.1	45.2	0.989	0.0107	1.000
A4r2	DEX.VF3	20	-5.0	10.1	51.2	0.850	0.1628	0.403
A4r2	SAD.VF3	20	-6.0	12.1	53.3	0.989	0.0111	1.005
A4r2	DEX.VF5	22	-3.3	10.6	63.9	0.912	0.0919	0.539
A4r2	LP.R4r2	21	-1.2	4.5	51.7			

^A Bold figures denote 'best' model ^B Number of parameters

correlations of unity, i.e. the CF effectively reduced to a VF. For stationary RF, p.e. correlations followed a smooth, steadily declining curve. In contrast, estimates for LP increased sharply for lags greater than 80 days, reflecting an erratic peak in estimates of correlations between youngest and oldest ages, previously attributed to problems due to small numbers of observations (Meyer 1999).

environmental variation than for the CF. Conversely, for $e=7$ LP models performed better than RF models with the same number of parameters, as the multiple measurement error variances had already largely accounted for variance heterogeneity. BIC involved a more stringent penalty for the number of parameters estimated and thus selected more parsimonious models as the 'best' model to describe the data. Whilst AIC favoured model A4r2.LP.R4r2 with $e=7$ and 21 parameters, i.e. a cubic regression on LP for both genetic and p.e. effects as selected in previous analyses (Meyer 1999), BIC considered model A3r1.DEX.VF5 for $e=1$ with 12 parameters adequate. Estimates of variance components for these and selected other models are shown in Figure 2. Overall, there was little difference in estimates of the total, phenotypic variation, even for an LP model with $e=1$. Partitioning of individual components varied, however, with noticeably more environmental variation 'picked up' by the p.e. VF for parametric correlation models than for the LP models, in particular for $e=7$. Corresponding average correlations for each lag are shown in Figure 3. Again, there was little difference between the models shown in estimates of phenotypic correlations. A genetic CF with rank 1 implied genetic

CONCLUSIONS

RR analyses modelling p.e. covariances between repeated records through a parametric CF and polynomial VF provide a parsimonious alternative to models fitting a regression on LP. Imposing a structure on p.e. correlations can alleviate problems of erratic estimates of covariances between oldest and youngest ages.

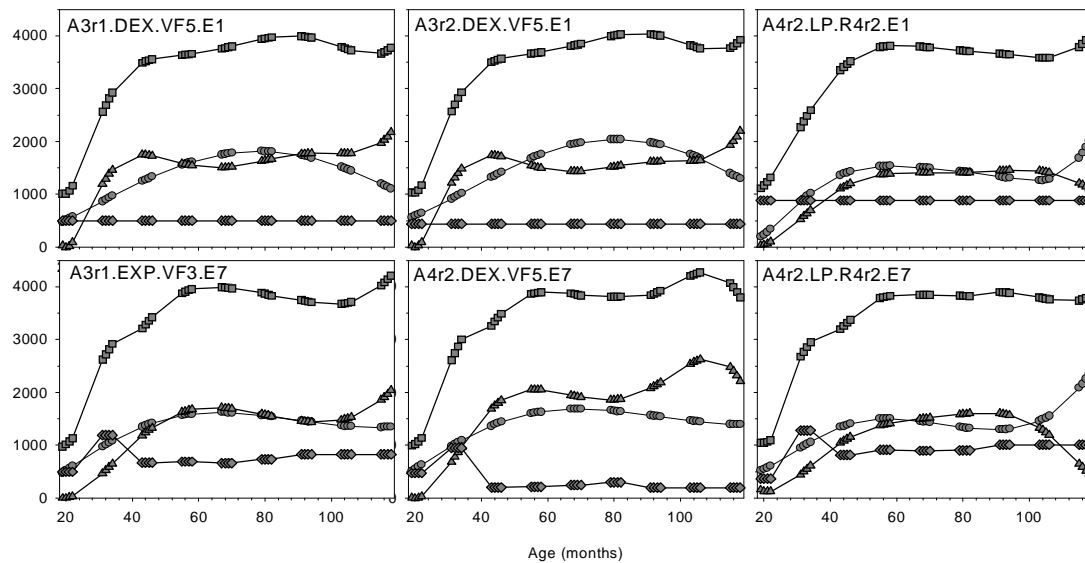


Figure 2. Estimates of variance components (■ : phenotypic, ● : genetic, ▲ : permanent environmental, and ◆ : measurement error).

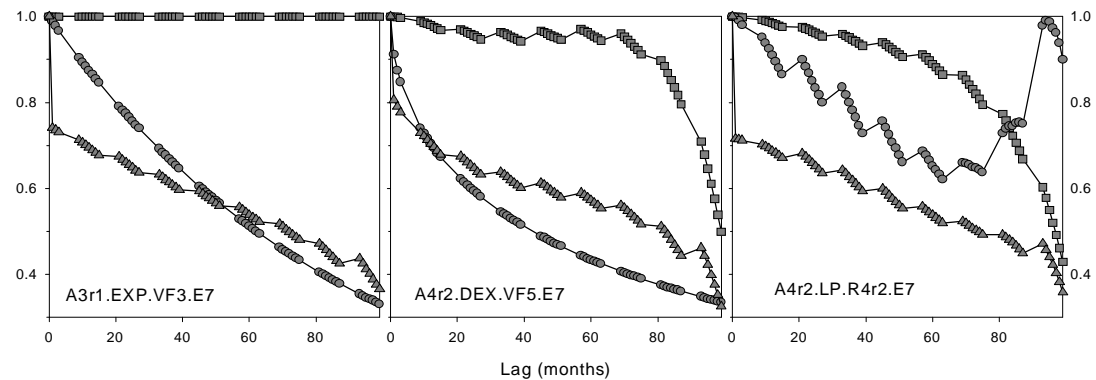


Figure 3. Estimates of average correlations for different lags (▲ : phenotypic, ■ : genetic, and ● : permanent environmental).

REFERENCES

- Meyer, K. (1999) *J. Anim. Breed. Genet.* **116** : 181.
 Nuñez-Antón, V. and Zimmerman, D.L. (2000) *Biometrics* **56** : 699.