ESTIMATES OF COVARIANCE FUNCTIONS FOR GROWTH OF AUSTRALIAN BEEF CATTLE FROM A LARGE SET OF FIELD DATA

K. Meyer

Animal Genetics and Breeding Unit, University of New England (jointly with NSW
Agriculture), Armidale, NSW 2351, Australia

INTRODUCTION

Random regression (RR) models for longitudinal data often require a considerable number of parameters to describe the associated covariance functions. Estimation of these parameters has proven cumbersome and subject to sampling problems. Moreover, computational restraints have limited analyses of data sets sufficiently large to supply accurate estimates. This holds especially for covariance functions pertaining to early growth of beef cattle, which is subject to both genetic and permanent environmental maternal effects. Analyses so far used restricted maximum likelihood (REML) estimation, and considered less than 4000 animals with records (**??**). Bayesian analysis provides an alternative which is simpler to implement and requires less computational resources per sample than REML, thus facilitating larger scale analyses. This paper presents a RR analysis of a large set of weight records from commercial beef cattle, considering weights from birth to 820 days of age. A Gibbs sampling algorithm is used to estimate covariance functions for direct and maternal effects.

MATERIAL AND METHODS

Data. Data consisted of weight records for Australian and New Zealand Hereford and Polled Hereford cattle. Raw data comprised more than 1.44 million weights, and pedigree information for 0.92 million animals. After initial edits, means and standard deviations for ages at recording, grouped in 7-day intervals were calculated. Weights up to 820 days of age were extracted for all animals in 90 herds, which had at least three animals with 6 or more records or in which least 40% of animals had 4 or more observations, disregarding any observations more than 3 SD from the respective, weekly mean. Further edits eliminated records clearly ’out of sequence’. After inspection of daily growth rates in the complete data set, any records taken more than 45 days apart and corresponding to a daily growth rate less than -1 or greater than 2.5 kg/day since the previous weighing were disregarded.

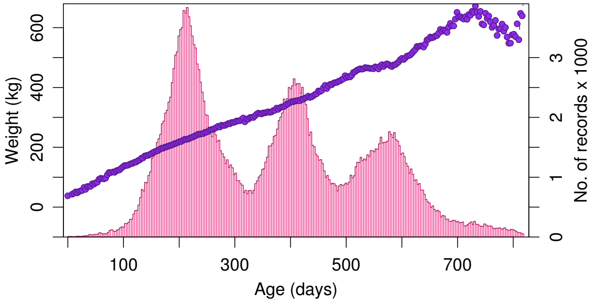

Figure 1. Numbers of records (■) and means (●) for ages at recording (in 3 day intervals)

Data structure. This yielded a total of 367,942 records on 130,412 animals, with a mean weight of 277.9 kg and mean age of 298.6 days. Birth weight was recorded for 73,360 animals. Figure 1 shows the distribution of weights over other ages at recording, together with corresponding means. Ages below 80 or above 700 days were grouped in 3 day intervals to increase the number of records per age. There were 25,242, 31,474, 33,365, 29,299 and 8,070 animals with 1, 2, 3, 4 and 5 records, respectively, and 2,962 animals had 6 or more weights recorded, up to a maximum of 14. Animals with records were progeny of 4,776 sires and 43,706 dams. Pedigree information for parents not in the data was obtained for up to 4 generations backwards, resulting in 176,718 animals in the analysis. Whilst only 31% of animals in the data had 4 or more records themselves, 21,204 and 49,308 of the remainder had a parent or a sib, respectively, with at least 4 records, i.e. at least 85% of animals in the data had sufficient information available to allow a cubic polynomial to be fitted.

Analyses. Estimates of (co)variance components were obtained by Bayesian analysis, employing a Gibbs sampling algorithm. Details are analogous to those for RR analyses of test-day records in dairy cattle; see **?** or **?**. Four Markov chains with a total of 390,000 samples were generated, assuming flat priors. Variances and genetic parameters for ages in 10 day intervals were calculated for each sample. Estimates were obtained as posterior means, disregarding the first 15,000 samples in each chain as "burn-in" period.

Random effects. The model of analysis fitted a cubic RR (4 coefficients) on Legendre polynomials (LP) of age at recording for direct (A) and maternal (M) genetic effects, and direct (R) and maternal (C) permanent environmental (PE) effects. Residual effects were considered independently distributed with heterogeneous measurement error variances (σ_i^2). Changes in σ_i^2 with age were modelled as a step function with 20 classes (0, 1–30, …, 271–300, 301–360, …, 721–780 and 781–820 days). This gave a total of 60 covariance components to be estimated.

Fixed effects. Changes in mean with age were modelled through fixed, cubic regressions on LP of age, nested within sex, dam age class (in years, treating ages $>$ 9 years as one class) and birth type (single *vs.* twin). In addition, contemporary groups (CG) were fitted as cross-classified fixed effects (43,605 levels). CG were defined as herd-sex-management group-year/month of weighing subclasses for birth weights, and herd-sex-management group-date of weighing subclasses otherwise. To reduce the range of ages compared directly, CG classes were further subdivided applying an "age slicing" of 45 days up to 300 days, and 60 days for higher ages.

Computing requirements. This yielded a mixed model with more than 2.17 million equations and 55.6 million non-zero off-diagonal elements in the lower triangle of the corresponding coefficient matrix. Computational demands were substantial, requiring about 19 seconds CPU time per sample on a Compaq Alpha ES45 machine with a 1 Ghz processor, and 873 Mbytes of memory.

Table 1. Estimates of covariances among random regression coefficients for genetic (upper triangle) and permanent environmental (lower triangle) effects

	Direct effects					Maternal effects				
	Intercept		Linear	Quad.	Cubic	Intercept		Linear	Quad.	Cubic
Intercept.	776.11	607.96	353.57	51.47	20.66	274.52	121.01	34.36	-2.83	18.49
Linear	338.55	281.31	244.85	71.37	22.53	36.85	9.94	16.87	2.97	3.48
Quadratic	-34.00	79.78	138.09	69.12	28.63	-16.85	2.25	8.13	5.79	1.64
Cubic	31.28	8.48	51.52	49.11	16.38	61.15	8.48	-0.81	16.19	5.97

RESULTS AND DISCUSSION

Estimated covariance matrices for the four sets of RR coefficients are given in Table 1. As observed previously, correlations between intercept and linear RR coefficients were high. Both maternal effects also showed a strong association between intercept and cubic term. This resulted in one eigenvalue amounting to less than 0.1% of their sum, suggesting that a reduced rank fit may suffice.

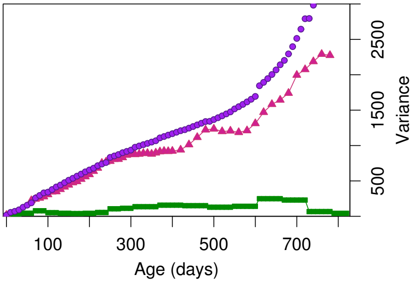

Figure 2. Phenotypic (● random regression, ▲ univariate analyses) and error (🟩) variances (kg^2)

Variances. Figure 2 shows phenotypic variances (σ^2_P) for the ages in the data together with corresponding univariate (UV) estimates from analyses of successive, partially overlapping data sets obtained by **?**, and estimates of σ^2_i for the 20 age classes. RR estimates of σ^2_P were consistently higher than those from UV analyses, increasingly so after 300 days of age. The latter could, in part at least, reflect effects of selection bias on UV estimates due to culling after weaning. Furthermore, UV analyses considered age ranges of 45 days, i.e. RR with an age slicing of 60 days after 300 days of age compared weights over a greater range of ages directly, which might have contributed to higher values for σ^2_P. After 700 days, RR estimates of σ^2_P increased sharply to an implausible value of almost 4200 kg^2 at 800 days, mainly due to a sharp increase in estimates of the direct PE variance (not shown). Similar, erratic behaviour of covariance function estimates for the highest ages with least records has been observed repeatedly, and remains an unresolved problem.

Genetic parameters. Estimates of genetic parameters are given in Figure 3, together with their approximate 95% highest posterior density regions (dashed lines). RR estimates for birth weight were 0.40, 0.07 and 0.05 for direct (h^2) and maternal (m^2) heritabilities and the proportion of variance due to maternal PE effects (c^2), respectively, and agreed closely with UV results. As observed previously (**??**), h^2 decreased sharply after birth and increased again after about 100 days of age, to 0.15, 0.24 and 0.36 at 200, 400 and 600 days, respectively. RR estimates of m^2 showed a peak at 160 – 200 days of age similar to that reported by **?**. Whilst the estimate of 0.09 for m^2 at 200 days was lower than UV results, it was consistent with m^2 values for weaning

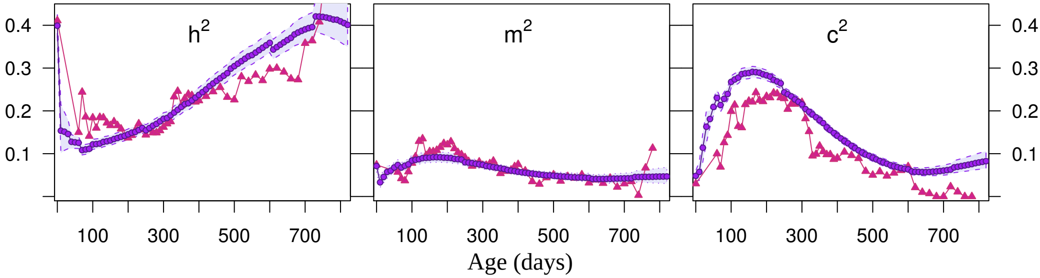

Figure 3. Genetic parameters from random regression (●) and univariate (▲) analyses (h^2 direct and m^2 maternal heritability, c^2 permanent environmental maternal effects)

weight, obtained considering records from 80 to 300 days. Conversely, estimates of c^2 of 0.29 at 160 and 0.28 at 200 days were higher than UV figures. With an age range of only 45 days in UV analyses, there was a relative paucity of records for dam-offspring pairs and multiple progeny per dam, i.e. UV estimates were likely to be subject to problems in partitioning maternal effects into their genetic and PE components. Estimates of both maternal variances remained virtually constant from 300 to 600 days. Hence m^2 and c^2 declined gradually with age, reflecting the diminishing importance of 'carry-over' maternal effects. Estimates of maternal PE variances and c^2, however, increased again implausibly after 700 days. Whilst analyses showed some maternal effects up to about 700 days, genetic evaluation schemes often consider these to be absent after 500 days of age. More sensible estimates of covariance functions for M and C might be obtained restricting their range.

Correlations. Figure 4 shows contour lines for correlations among ages in the data, in steps of 0.05 plotted against ages in days. Both genetic (r_A) and PE (r_R) direct correlations decreased steadily with increasing lag in age at recording. Estimates agreed with literature results which generally report a higher r_A between birth and final than with intermediate weights. Maternal genetic (r_M) and PE (r_C) correlations formed a plateau with correlations largely higher than 0.9 (r_C) and 0.8 (r_M) and, except for the youngest ages, few values below 0.7.

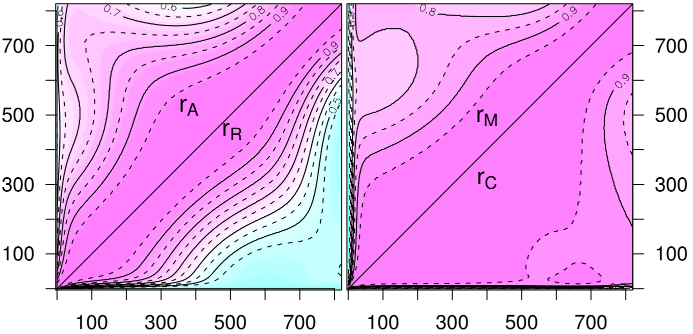

Figure 4. Direct (r_A) and maternal (r_M) genetic, direct (r_R) and maternal (r_C) permanent environmental correlations

CONCLUSIONS

Overall estimates showed reasonable agreement with previous results and corresponding univariate analyses, suggesting that changes in covariances between weights of beef cattle with age can be modelled adequately through random regressions. In spite of a large data set, problems with erratic estimates at the highest ages were apparent. Modelling the 'plateau' of maternal correlations close to unity through a cubic polynomial yielded inflated maternal variances at higher ages. Further work should examine different orders of fit, models restricting maternal effects, and the scope for reduced rank models.

REFERENCES

Albuquerque, L. G. and Meyer, K. (2001) *J. Anim. Sci.* **79** : 2776–2789.

Jamrozik, J. and Schaeffer, L. R. (1997) *J. Dairy Sci.* **80** : 762–770.

Meyer, K. (2001a) *Genet. Sel. Evol.* **33** : 487–514.

Meyer, K. (2001b) *Proc. 14th Conf. Assoc. Advanc. Anim. Breed. Genet.* Queenstown, NZ, pp. 147–150.

Rekaya, R., Carabaño, M. J. and Toro, M. A. (1999) *Livest. Prod. Sci.* **57** : 203–217.