Estimates of genetic covariance functions for growth of Australian Angus cattle from random regression models fitting different orders of polynomials

Karin Meyer

Animal Genetics and Breeding Unit[1], University of New England, Armidale NSW 2351, Australia

INTRODUCTION

Random regression (RR) analyses have been suggested for the analysis of growth records for beef cattle. A previous analysis of a large set of field data fitted cubic regressions on age at recording for all random effects, i.e. direct and maternal, genetic and permanent environmental effects (Meyer, 2002a). This choice was determined by the fact that a cubic polynomial is the lowest degree polynomial for which both first and second derivatives, i.e. growth rate and acceleration, vary with age, and it involved similar numbers of genetic coefficients than current multi-trait analyses of birth, weaning, yearling and final weights. In addition, few animals had more than 4 weight records, limiting the scope for estimating higher order polynomials. Results from restricted maximum likelihood analyses of small data sets with more records per animal available, suggested that models with differential orders of fit for individual random effects, would be advantageous (Albuquerque and Meyer, 2001a; Meyer, 2001).

This paper presents examines the effects of different orders of polynomial fit on estimates of covariance functions for growth of Australian Angus cattle, considering a moderately large set of field data, selected to have a high portion of animals with 4 or more records using Bayesian analysis.

MATERIAL AND METHODS

Data. Data comprised records for weights of Australian Angus cattle, taken from birth to 820 days of age. After basic edits, records for animals in 43 herds with an average number of records per animal greater than 3.3 or at least 50% animals with 4 or more records were extracted from the National Beef Recording Scheme data base. Based on the inspection of growth rates, further edits eliminated any records clearly 'out-of sequence' for individual animals. Recursively deleting any records on animals with less than 3 observations and any records in single record subclasses, then left $84,533$ records on $20,731$ animals. Figure 1 shows the distribution of records over ages at recording in 5-day intervals (except for $17,891$ birth weights), together with corresponding mean weights. With a maximum of 9 and an average of 4.08 records per animal, 70.9%, 27.7% and 7.3% of animals had at least 4, 5, and 6 weights recorded. Considering information from parents, progeny and sibs, corresponding proportions were 98.3%, 85.2% and 48.1%, respectively.

Pedigree information up to 5 generations backwards was considered. Pedigrees were 'pruned', i.e. any parents without records and a link to a single offspring only were treated as unknown. This was done recursively, resulting in $14,326$ parents without records to be included in the analysis, i.e. $35,057$ animals in total. After pruning, there were 1345 sires and 7494 (7977 before pruning) dams with progeny in the data. Pruning maternal genetic effects separately yielded $20,463$ effects to be fitted. With a considerable amount of embryo transfer animals in the data, there were 9670 permanent environmental dams.

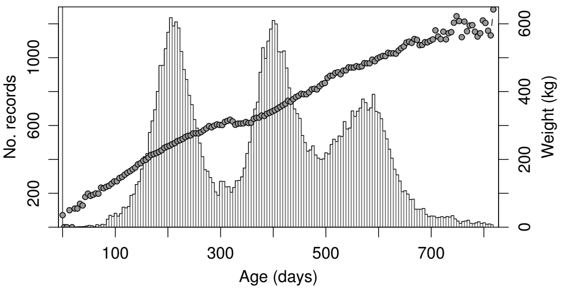

Figure 1. Numbers of records (■) and means weights(●) for ages at recording in 5-day intervals

Analyses. Estimates of (co)variance components were obtained by Bayesian analysis, employing a Gibbs sampling algorithm as implemented in program RRGIBBS (Meyer, 2002b). Three Markov chains with $100,000$ samples each were drawn, assuming flat priors. Variances and genetic parameters

---

[1] a joint venture with NSW Agriculture

for ages in 10 day intervals were calculated for each sample. Estimates were obtained as posterior means, disregarding the first $20,000$ samples in each chain as "burn-in" period. Solutions for all effects fitted were obtained by Best Linear Unbiased Prediction for the estimates of covariance components for each model, and residuals and their sums of squares (SS) were calculated.

Random effects. The model of analysis fitted a RR on Legendre polynomials (LP) of age at recording for four random effect, considering various orders of ft (k). Direct, additive genetic effects (A) were, in turn, modelled through quadratic, cubic and quartic polynomials, involving $k_A = 3,4$ and 5 RR coefficients, respectively. Direct permanent environmental effects (R) were modelled analogously, considering a quintic polynomial in addition, i.e. $k_R = 3,4,5,6$. Maternal genetic (M) and permanent environmental (C) effects were modelled through either quadratic or cubic polynomials, i.e. $k_M = k_C = 3,4$. To avoid problems with spurious estimates at the highest ages, their effects were restricted to ages from 0 to 600 days, except for one analysis. Direct and maternal genetic effects were assumed to be uncorrelated. Models fitted are denoted in the following by the number of RR coefficients for A, M, R and C, respectively. For example, M4353 describes an analysis fitting quadratic polynomials for both maternal effects, and cubic and quartic polynomials for direct genetic and permanent environmental effects, respectively. Combinations examined were M3333, M3343, M3353, M3363, M4343, M4444, M4353, M5353, M5454 and M5363. Residuals 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 19 classes (0, 1–30, …, 271–300, 301–360, …, 721–780 and 781–820 days) throughout. This gave a total of 43 (M3333) to 69 (M5454) 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 (8822 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 for other weights. 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. If this resulted in a small subclass with less than 5 records for the highest ages in the CG, this last subclass was merged with the previous age subclass, provided the range of ages did not exceed 54 days for weights up to 300 days and 72 days otherwise.

Univariate analyses. In addition, eight corresponding univariate analyses were carried out by restricted maximum likelihood, considering ages of $0, 1-200, 101-300, \ldots$ and $601-820$ days. These considered single records per animal only, fitting direct and maternal genetic effects and maternal permanent environmental effects. Fixed effects considered were CGs, as defined above, the effects of birth type and dam age class (heifer *vs.* cow). Dam age was fitted as a linear and quadratic covariable, and age at recording was accounted for by fitting a linear regression for each sex of animal.

RESULTS

Average residual SS together with the number of parameters (p) estimated and estimates of the first two eigenvalues (λ_1 and λ_2) and the sum of eigenvalues ($\sum\lambda$) of the covariance matrix for direct genetic RR coefficients are summarised in Table 1. Clearly, residual SS were dominated by the order of polynomial fitted for direct, permanent environmental effects, with only a slight decrease in SS with increasing values for k_A. The total variation attributed to direct genetic RR coefficients ($\sum\lambda$) was determined by k_A, with values highest for a cubic polynomial. λ_1 explained 92–95% of variation for all models, with λ_2 amounting to 4–7%. Figure 2 shows estimates of phenotypic (σ_P^2) and direct, permanent environmental (σ_R^2) variances. On the whole, estimates from all RR models agreed well for the ages with substantial numbers of observations. Some differences, more notable for the variance ratios shown below, were apparent for records after birth to about 150 days of age, and there were big discrepancies between estimates from about 700 days with a clear effect of the order of fit for R. In particular, estimates of σ_R^2 for $k_R = 4$ increased dramatically and implausibly for the highest ages. Values of σ_P^2 from univariate analyses agreed with their counterparts from RR analyses until weaning, but tended to be somewhat lower for higher ages. This could, in part at least, be due to effects of

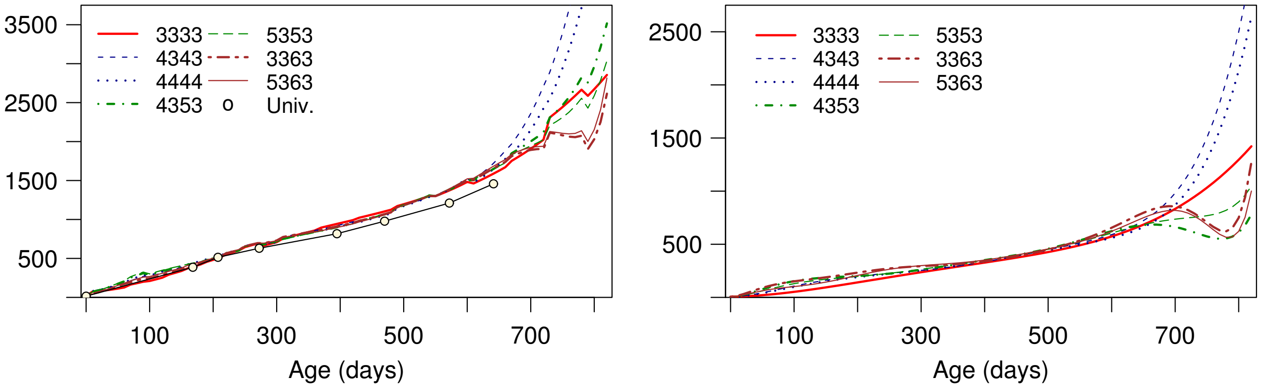

Figure 2. Estimates of phenotypic (left) and direct permanent environmental (right) variances.

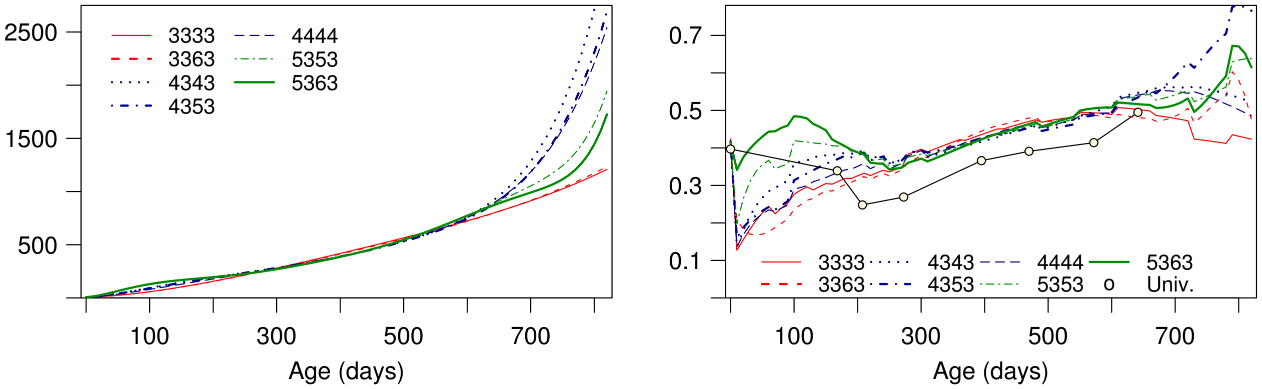

Figure 3. Estimates of direct genetic variances (left) and heritabilities (right).

selection on weaning weight on univariate estimates.

Similarly, estimates of direct genetic variances (σ^2_A) and heritabilities (h^2), shown in Figure 3, for the highest and youngest (except birth weight) ages were in the main determined by k_A. Some effect of the order of fit for the other random effects could be noted, however, emphasizing the partitioning of total variation according to the model assumed for all sources of variation. Estimates of h^2 for birth weight from all analyses were consistently around 40%, the value expected from literature results, with narrow confidence regions. RR estimates from 200 to about 650 days were similar for all models. Whilst they showed the same trend as univariate estimates, they were consistently and substantially higher than the latter. There are few h^2 estimates for weights of beef cattle at early ages. Albuquerque and Meyer (2001b) reported a steep decline in h^2 after birth. Estimates from models with $k_A = 3$ were most consistent with such pattern.

Table 1. Residual SS and eigenvalues

Model	p	SS	λ_1	λ_2	$\sum\lambda$
M3333	43	53.6	911.2	42.1	956.3
M3343	47	36.91	915.5	47.4	963.7
M4343	51	36.4	1183.0	91.3	1281.4
M3353	52	30.1	914.4	46.1	963.5
M4353	56	29.6	1112.6	71.0	1190.0
M3363	58	24.7	905.5	46.1	954.6
M4444	59	35.4	1098.1	72.3	1177.4
M5353	61	29.6	994.5	59.1	1065.5
M5363	67	24.5	938.4	68.6	1019.0
M5454	69	28.7	1010.3	45.9	1068.2

Figure 4 shows estimates of maternal heritability (m^2) and the proportion of variance explained by maternal permanent environmental effects (c^2). Again, there were differences in estimates determined by the order of fit, especially for the youngest ages with few records. Not restraining the age range for maternal effects resulted in spurious estimated of maternal effects for the highest ages. Values for m^2 were consistently lower than results from univariate analyses, and somewhat lower than expected for Angus cattle. For instance, Meyer (1997) reported estimates of 8 to 12 % for m^2 for weaning weight of Australian and New Zealand Angus, whilst our estimates for 200 day weight from RR analyses ranged from 4 to 6%. This suggested that the higher estimates of σ^2_A were, in part at least, due to a differential partitioning of direct and genetic maternal effects between types of analyses.

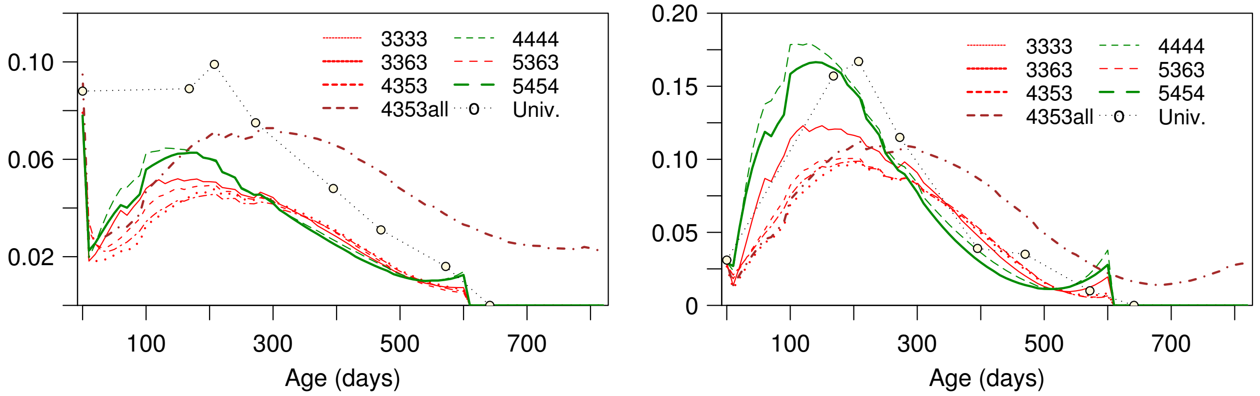

Figure 4. Estimates of maternal heritability (left) and permanent environmental variance ratio (right).

DISCUSSION

Results do not provide an unequivocal answer as to which model might be "best" for RR analyses of weight records in beef cattle. Clearly, results from RR analyses are only as good as the data they are based on. If there are few records for some ages, estimates of variances and genetic parameters for these ages are likely to be determined by the shapes of polynomials fitted, and can be misleading.

Data utilised in this study had as good a structure as might be expected for field records on beef cattle. In spite of this, fitting regressions on higher order polynomials proved problematic. The previously favoured cubic polynomial was found to give rise to most end-of-range problems, yielding implausible estimates of variances for the highest ages. In the absence of more frequent weight recording, a model fitting quadratic regressions for all random effects (M3333) might be the conservative and sensible choice. Figure 5 shows estimates of genetic parameters for this model.

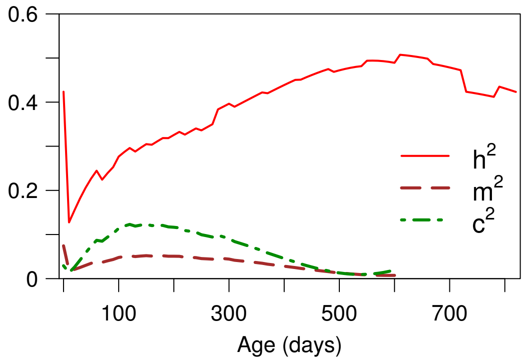

Figure 5. Estimates for model M3333

Estimates of covariance matrices among RR coefficients had at least one eigenvalue close to zero, suggesting that reduced rank estimates might be adequate and provide more stable estimates. Ultimately, we want to model lifetime growth of beef cattle in a RR framework. For a large range of ages, a single low-order polynomial might not be adequate, and we might want to consider a segmented curve, i.e. splines. Analyses fitting a RR on linear and quadratic B-spline functions of age are in progress for the data in this study.

ACKNOWLEDGMENTS

This work was supported by Meat and Livestock Australia Ltd. under grant BFGEN.100.

REFERENCES

Albuquerque L.G. and Meyer K., 2001a. Estimates of covariance functions for growth from birth to 630 days of age in Nelore cattle. J. Anim. Sci. 79:2776–2789

Albuquerque L.G. and Meyer K., 2001b. Estimates of direct and maternal genetic effects for weights from birth to 600 days of age in Nelore cattle. J. Anim. Breed. Genet. 118:83–92

Meyer K., 1997. Estimates of genetic parameters for weaning weight of beef cattle accounting for direct-maternal covariances. Livest. Prod. Sci. 52:187–199

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

Meyer K., 2002a. Estimates of covariance functions for growth of Australian beef cattle from a large set of field data. CD-ROM Seventh World Congr. Genet. Appl. Livest. Prod. Communication No. 11–01

Meyer K., 2002b. "RRGIBBS" – A program for simple random regression analyses via Gibbs sampling. CD-ROM Seventh World Congr. Genet. Appl. Livest. Prod. Communication No. 28–27