

THE EFFECT OF INCORRECT ESTIMATED VARIANCE-COVARIANCE COMPONENTS ON GENETIC EVALUATION OF DAIRY CATTLE WITH RANDOM REGRESSION MODELS

T. Strabel¹ and J. Jamrozik²

¹Department of Genetics and Animal Breeding, Agricultural University of Poznan, ul. Wolynska 33, 60-637 Poznan, Poland

²Centre for Genetic Improvement of Livestock, Department of Animal and Poultry Science, University of Guelph, Guelph, ON, Canada, N1G 2W1

INTRODUCTION

Random regression test-day models (RRM) are preferred above lactation models for milk production traits because of their abilities to model a separate lactation curve for every animal and to account for short-term environmental effects. Precise estimates of genetic parameters are needed for genetic evaluation with models using test-day data. Estimating variance components of a RRM is not a straightforward task. Misztal *et al.* (2000) showed that the range and pattern of variances and heritability across days in milk vary widely among studies on RRM. Differences in estimated parameters may be due to the size of the data and the functions used to describe random regressions. Third and fourth order Legendre polynomials were suggested (Pool and Meuwissen, 2000) as a compromise between model complexity and accuracy of genetic evaluation. If estimated parameters are incorrect, however, the advantage of RRM over lactation model may not be fully realized. The objective of this simulation study was to analyze the impact of incorrect variance components on the estimation of breeding values with RRM.

MATERIAL AND METHODS

Data consisted of 54,549 test-day (TD) milk yields of 6347 first lactations of Polish Black and White cows collected in 737 herds in Wielkopolska region. Cows were assigned to 25 subclasses for age-season of calving. The following model was used :

$$Y_{ijklm} = \text{HTD}_i + \sum b_{jl} z_{mnl} + \sum c_{kl} z_{mnl} + \sum a_{ml} z_{mnl} + \sum p_{ml} z_{mnl} + e_{ijklm} \quad (1)$$

where Y_{ijklm} is the l -th TD milk yield of cow m within herd-test day effect i , for a cow from herd k belonging to subclass j of age-season of calving, HTD_i is random herd-test day effect, b_{jl} are fixed regression coefficients specific to age-season subclass j , c_{kl} are fixed regression coefficients specific to herd k , a_{ml} are random regression coefficients specific to cow m , p_{ml} is random permanent environmental (PE) effect, e_{ijklm} is residual effect for each observation, z_{mnl} are covariates. The third order Legendre polynomials were used as covariates for both fixed and random regressions. Residual variance was constant in the interval from 5 to 305 days in milk (DIM) and all random effects of the model were assumed to be normally distributed.

Variance components (VC) were obtained for all random effects of the model using the REML method and the estimates were named TRUE VC. Utilizing data structure as described above, TRUE VC and solutions of fixed effects from the REML analysis, 60 new data sets were generated using model (1). Variance components were estimated for each set of simulated data. Four 'extreme' cases of components were selected (VC1 – VC4) based on the shapes of daily heritability and PE variance ratio for further use. TRUE VC and solutions for fixed effects were used again to simulate 40 sets of true breeding values (BV) and phenotypic observations. Each of these data sets was analyzed by model (1) with 14 different sets of VC, which were as follow: TRUE VC, 4 selected earlier (VC1-VC4), 2 sets of VC (VC5 and VC6) estimated for the same population with one and two-lactations models, respectively (Strabel and Misztal, 1999), and finally, the set of 7 VC (VC7-VC13), where genetic and permanent variances were constant across all DIM. In the VC7 set the heritability and the ratio of PE to the total variance were equal to the average corresponding values for the set of TRUE VC. Sets VC8-VC13 were obtained by modification of VC7, where various combinations of higher/lower (+/-20%) heritability and/or repeatability were assumed. The BV and estimated breeding values (EBV) for the total milk yield in lactation were compared by calculating bias (defined as BV-EBV), mean squared difference and product moment correlation. Results were averaged over 40 replicates.

RESULTS AND DISCUSSION

Daily heritability and PE to total variance ratio for the TRUE VC are shown in Figures 1 and 2, respectively. The range of values for these parameters are similar to those obtained by Strabel and Misztal (1999) for the same population. Heritabilities at the beginning and the end of lactation were smaller than the values for mid-lactation DIM. An additional set of fixed regressions was used in this study which accounted for the differences in shapes of lactation in various herds.

Table 1 presents mean bias, square root of mean squared differences (MSD) and correlations between BV and EBV for bulls with more than 30 daughters and for cows with TD records. Average difference between BV and EBV estimated by model with TRUE VC was equal to – 5.4 for bulls and –0.1 for cows. Bias of EBV from models using other sets of VC indicate that the use of incorrect VC had practically no influence on EBV. This is in agreement with Henderson (1975). All sets of VC caused a small increase in MSD between true and estimated BV for bulls and cows. The set of VC estimated for the same population and slightly different model (VC6) and the set that overestimated genetic variance in the middle and the end of lactation (VC2) gave slightly lower values of MSD. Mean correlations between TRUE and estimated breeding values were larger than 0.82 for bulls with more than 30 daughters and 0.54 for cows with TD records. Correlations were slightly lower for all sets of VC when compared to the TRUE VC. Models VC7 – VC13 were in fact 'fixed regression' models where random regressions were reduced to a constant function over all DIM. This indicate that the breeding value for the total yield in lactation can be predicted accurately enough by using simplified RRM.

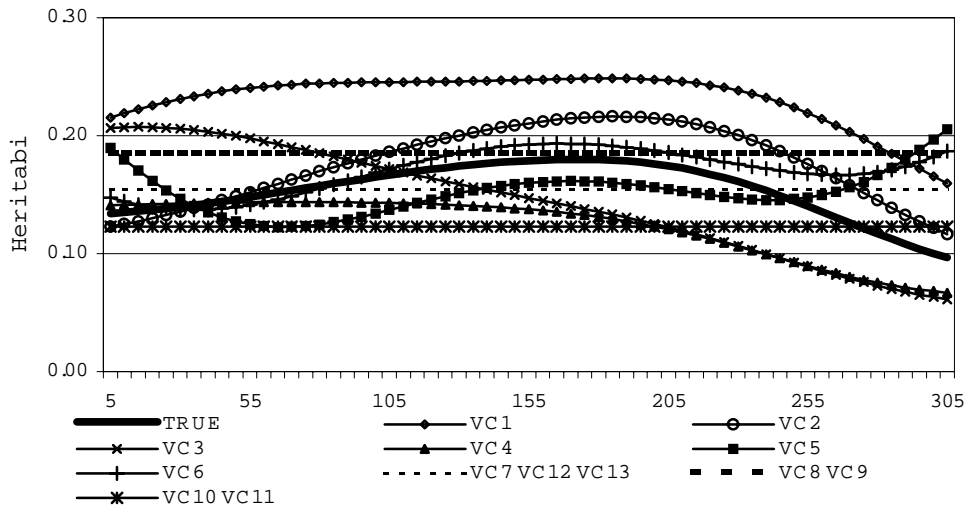


Figure 1. Heritability of daily milk yield for models with TRUE and selected sets of variance components

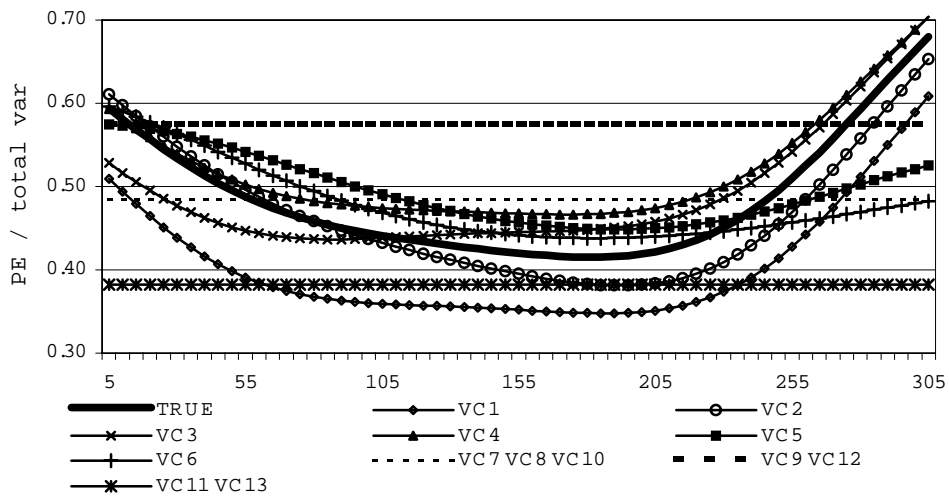


Figure 2. Ratio of daily permanent environmental variance to the total variance for TRUE and selected sets of variance components

Table 1. Mean bias, square root of mean squared differences (MSD) and correlations between BV and EBV for bulls with more than 30 daughters and for cows with lactation records

Model	Bias		MSD		Correlations	
	Bulls	Cows	Bulls	Cows	Bulls	Cows
TRUE VC	-5.4	-0.1	163.5	252.8	0.836	0.560
VC1	-7.0	-0.5	167.1	263.3	0.834	0.549
VC2	-6.3	-0.3	164.5	254.2	0.835	0.557
VC3	-4.0	0.2	167.2	256.6	0.829	0.542
VC4	-4.0	0.2	165.7	255.3	0.835	0.554
VC5	-4.5	0.2	165.5	255.3	0.835	0.555
VC6	-5.4	-0.1	164.0	253.5	0.835	0.558
VC7	-4.8	0.1	165.7	254.5	0.832	0.554
VC8	-5.5	0.1	165.9	254.2	0.832	0.554
VC9	-5.5	0.1	166.5	254.5	0.830	0.553
VC10	-4.0	0.3	167.4	256.7	0.832	0.550
VC11	-4.8	0.1	165.2	254.1	0.833	0.554
VC12	-4.9	0.2	166.8	255.4	0.831	0.552
VC13	-5.1	0.1	165.3	254.1	0.833	0.554

The effect of incorrect estimated parameters on BV estimation was investigated by Schaeffer (1975). He found that the use of biased VC increases the prediction error variance (PEV) which is in agreement with results obtained in this study. Schaeffer concluded that traits with smaller heritability suffer greater increase of PEV. This would suggest, that small impact of the incorrect VC on BV found in this study could be even lower for populations with higher relative genetic variance. Several assumptions were made in this study. Parameter estimation, data simulation and breeding value prediction were all done using the same model. Data simulation process was based on a real data structure and it did not allow for selection. Only one trait (e.g. total yield) was analyzed. The impact of incorrect estimated VC may be more significant on individual regression coefficients or on the measure of lactation persistency. Further studies including multiple-trait RRM are warranted.

REFERENCES

- Henderson, C.R. (1975) *J. Anim. Sci.* **41** : 760-770.
 Jamrozik J. and Schaeffer, L.R. (1997) *J. Dairy Sci.* **80** : 762-770.
 Misztal, I., Strabel, T., Jamrozik, J., Mantysaari, E., and Meuwissen, T. (2000) *J. Dairy Sci.* **83** : 1125-1134.
 Pool, M.H. and Meuwissen, T.H.E. (2000) *Livest. Prod. Sci.* **64** : 133-145.
 Schaeffer, L.R. (1984) *J. Dairy Sci.* **67** : 1567-1580.
 Strabel, T. and Misztal, I. (1999) *J. Dairy Sci.* **82** : 2805-2810.