

A SIMPLE BOOTSTRAPPING PROCEDURE TO VALIDATE THE MERINOSELECT MODEL FOR WEANING WEIGHT

A.A. Swan and D.J. Brown

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

SUMMARY

Sampling distributions for the regression of progeny performance on sires' estimated breeding values (EBV) were estimated for weaning weight in the MERINOSELECT database. Mean regression coefficients from these distributions were used to compare the effects of data quality and the model used for evaluation. Results showed that better quality data with dam pedigree and date of birth recorded, and an evaluation model which included sire by flock-year interaction, improved the prediction of progeny performance from estimated breeding values. Sampling was conducted at three different strata, the most effective division being at the level of contemporary groups.

INTRODUCTION

Sheep Genetics Australia (SGA) is a new across flock evaluation system for the Australian sheep industry launched in 2005. Within the SGA system, breeding values are estimated for Merino sheep in routine analyses which are separate from other breeds, and marketed under the MERINOSELECT brand name. Prior to the launch of SGA, an extensive research program was undertaken, heavily focused on developing data processing procedures and statistical models for the MERINOSELECT analysis (Brown *et al.* 2007). One of the methods used to validate models for each trait was to predict performance by estimating the regression of offspring records on their sires' estimated breeding values. Mostly these regression estimates were close to the expected value of 0.5, but for some traits they were significantly lower, at least on first inspection. This was particularly the case for traits measured early in an animal's lifetime. In this paper we investigate one of those traits, weaning weight, exploring in detail the impact of data quality and evaluation model on the prediction of progeny performance.

MATERIALS AND METHODS

Regression of offspring performance on sire EBV. Genetic evaluation models are often validated by estimating the regression of offspring performance on sire breeding value (sire-offspring regression), which has an expected value of 0.5 assuming random mating. The method involves splitting the complete data into two independent sets of progeny, estimating breeding values in the first, and then regressing offspring performance on the EBV of link sires in the second. Link sires in this context are sires that have progeny in both data sets.

Bootstrapping was used to estimate the sampling distribution of the regression coefficient by replicating the division of animals into the two data sets. The procedure involved: 1) taking a random permutation of animals and allocating the first half to the evaluation data set and the second half to the regression data set; 2) estimating breeding values in the evaluation data set; 3) fitting a fixed effects model to the regression data set, adjusting for the effects of contemporary group, birth and

* AGBU is a joint venture of NSW Department of Primary Industries and The University of New England

Genetic Evaluation

rearing type, age of dam, and animal age; 4) selecting the residual values from step 3 for the progeny of link sires and fitting a model to these including contemporary group as a factor to account for any sub-sampling effects in forming the regression data set, and sires' EBV as a covariate; and 5) replicating steps 1 to 4 fifty times. All analyses were performed using the statistical software ASReml (Gilmour *et al.* 2006).

The division of animals into data sets was stratified at one of three levels: by site (effectively flock of origin), by contemporary group, or by animal. For example, when sampling by site, we take a random permutation of site codes, and then allocate animals from sites in the first half of the permutation to the evaluation data set, and animals from sites in the second half of the permutation to the regression data set.

Data subsets. In order to test the impact of data quality the bootstrapping procedure was applied to three data sets, with the first consisting of all animals in the MERINOSELECT data base with weaning weight records (data set 1). The second (data set 2) was constructed by selecting animals from contemporary groups where there was variation in the age at measurement, implying that actual dates of birth had been recorded. The third (data set 3) was restricted to animals from contemporary groups with age variation, and also from sites where more than 75% of animals had known dams. This last set represents the highest level of data quality tested. A summary of the three data sets is shown in Table 1. Note that data sets 2 and 3 are subsets of data set 1.

Table 1. Summary of weaning weight data sets

Data set	Animals recorded	% with known:		Number of:		Number of sites	Contemporary groups:	
		sire	dam	sires	dams		number	av. size
1	209,601	79	50	3,267	56,543	140	2,043	103
2	106,432	86	73	2,111	41,583	86	1,154	92
3	57,809	93	95	1,339	28,209	42	704	82

Evaluation models. Two evaluation models were also compared, the first (model 1) approximating the standard MERINOSELECT analysis for weaning weight, and the second (model 2) including a sire by flock-year random effect in addition.

The routine MERINOSELECT analysis pre-adjusts data for the fixed effects of birth and rearing type, age of dam, and age of animal. The evaluation software then fits contemporary group as the sole fixed effect in an animal model with maternal effects. Genetic groups are also fitted as random effects, with animals allocated to groups by flock and time period (see Brown *et al.* 2007 for more details). In the ASReml analyses used in this paper, we fitted all fixed effects simultaneously (i.e. within the evaluation model), and did not fit genetic groups. In a further point of departure, the parameter estimates in our evaluation model were derived from analyses of data set 3 described above.

The impact of low quality data. In order to examine the effect of including lower quality data with data from well recorded groups, regression sampling distributions were derived firstly by performing evaluations in data set 1 (the complete data set), and then estimating sire-offspring regressions using only progeny from data set 3 (highest quality subset). By doing so, the sires in the regression are those which are present in well recorded groups, but their breeding values have been estimated from

groups of varying data quality. For these comparisons, all three sampling strata were used, but only one evaluation model was tested (model 2, with sire by flock-year interactions).

RESULTS

Parameter estimates from data set 3 are shown in Table 2, along with the corresponding values initially used in the MERINOSELECT evaluation. Compared to these initial parameters, the heritability estimated in the standard model is substantially higher, and the direct-maternal correlation is substantially more negative. Fitting the sire by flock-year interaction significantly improves the likelihood of the model, with the interaction variance ratio estimated at 0.04. The estimated heritability is lower (0.39 versus 0.49) and the direct-maternal correlation is less negative (-0.42 versus -0.53).

Table 2. Parameter estimates from data set 3 for weaning weight, including phenotypic variance (V_p), direct heritability (h^2), maternal heritability (m^2), direct-maternal genetic correlation (r_{am}), maternal permanent environment ratio (pe^2) and sire by flock-year variance ratio (s^2). The third line shows the parameters initially used in the routine MERINOSELECT evaluation

Model	V_p	h^2	m^2	r_{am}	pe^2	s^2
1	15.30 ± 0.12	0.49 ± 0.02	0.19 ± 0.02	-0.53 ± 0.03	0.05 ± 0.01	-
2	15.45 ± 0.13	0.39 ± 0.03	0.15 ± 0.02	-0.42 ± 0.04	0.05 ± 0.01	0.04 ± 0.00
MERINOSELECT	15.73	0.31	0.17	-0.39	0.05	-

Means and standard deviations from the sampling distributions of the sire-offspring regression coefficients are shown in Table 3. When the sampling strata was by animal, while there was little difference in the mean regression between data sets, fitting the interaction model increased the mean regression to approximately 0.60 such that progeny performance was over-predicted. For the other sampling strata, the mean regressions were lowest for data set 1 and highest for data set 3, with the interaction model again increasing the mean regression.

Table 3. Mean ± standard deviation for sampling distributions of sire-offspring regressions

Sampling strata	Data set 1		Data set 2		Data set 3	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Site	0.13 ± 0.04	0.18 ± 0.05	0.22 ± 0.09	0.30 ± 0.12	0.46 ± 0.16	0.61 ± 0.19
Contemporary group	0.38 ± 0.03	0.45 ± 0.03	0.39 ± 0.03	0.49 ± 0.04	0.41 ± 0.02	0.51 ± 0.03
Animal	0.54 ± 0.01	0.61 ± 0.01	0.48 ± 0.01	0.60 ± 0.01	0.48 ± 0.02	0.60 ± 0.02

In the analyses testing the impact of low quality data, the mean sire-offspring regression coefficients were 0.26 ± 0.11, 0.47 ± 0.03, and 0.56 ± 0.02 for the site, contemporary group, and animal sampling strata respectively.

DISCUSSION

The results demonstrate both the importance of fitting sire by flock-year interaction effects in the MERINOSELECT model for weaning weight, and the value of a comprehensive recording program.

The estimated sire by flock-year effects indicate that there are significant genotype by environment interactions in these data, although there may also be an influence from non-additive interactions with dam genotypes in a breed as diverse as the Merino. Significant sire by flock interactions have been observed consistently across a range of traits in the MERINOSELECT database, and at the time of writing were under consideration for inclusion in the evaluation model.

Predicted progeny performance improves when the animals in the evaluation have birth date, age at weaning, and dam pedigree recorded. This is especially important for traits such as weaning weight where birth and maternal effects account for a significant portion of the variation. Other traits which may be affected in the same way include body weights up to the hogget stage, and early fleece weights.

Lower quality data did influence the estimated breeding values of sires with progeny in well recorded groups, but mainly when stratification was by site. These regressions were close to the expected value for the other sampling strata. This indicates that lower quality data need not be excluded from the evaluation, provided sire interactions are fitted.

Stratification by contemporary group proved to be most reliable in these data. Stratification by site may have performed worst in the presence of genotype by environment interactions, because there was no overlapping information on interactions between the sites in the regression and evaluation data sets. That is, there were no common sites between data sets. Stratification by animals may have tended to over-predict performance because sire by flock-year effects were effectively removed from the comparisons, but they were still fitted in the evaluation model. For the purposes of prediction of progeny performance, the evaluation model may therefore have been over specified.

CONCLUSIONS

High quality recording and the inclusion of sire by flock-year interactions in the MERINOSELECT data improved the reliability of estimated breeding values. As a result of these findings, sire by flock-year interactions are under consideration for inclusion in the evaluation model, and the impact of data quality will be investigated for other traits across all SGA evaluations.

ACKNOWLEDGMENTS

This research was supported by the Australian sheep industry through the project SHGEN.113 funded by Meat Livestock Australia and Australian Wool Innovation Ltd.

REFERENCES

- Brown, D. J., Huisman, A.E., Swan, A.A., Graser, H.U., Banks, R.B., Ball, A.J., Atkins, K.D. and Woolaston, R.R. (2007) *Proc. Assoc. Advmt. Anim. Breed. Genet.* 17:187.
- Gilmour, A.R., Gogel, B.J., Cullis, B.R., and Thompson, R. (2006) ASReml User Guide Release 2.0 VSN International Ltd, Hemel Hempstead, HP1 1ES, UK