Adjustment for over-evaluation of young bulls based on parent average

D. L. Johnson[*] and B. L. Harris[*]

Introduction

Breeding values (BV) for dairy cattle are generally estimated using BLUP procedures and an animal model. Properties of the estimates, based on model assumptions, include minimum error prediction variance and unbiasedness. However, the estimated BV of young bulls based on parent average (PA) are typically higher than subsequent BV which include daughter information. Several studies have documented evidence of such slippage in estimated breeding merit (van der Werf et al. (1994); Pedersen et al. (1995); Zwald and Weigel (2002); Rensing et al. (2009)). This is generally not a problem while bulls are on lay-off awaiting progeny test, but is of concern when bulls are used widely at a young age, for example, using genomic information. This paper investigates possible causes of variation in the overestimation including bull's herd of origin, embryo transfer (ET) status and genetic merit of the bull's dam. The primary objective is to develop an adjustment method for the New Zealand Animal Evaluation system for dairy cattle so that young bulls being selected on genomic information are not systematically over-evaluated.

Material and methods

National BV, using a test day model, were generated at the end of each season commencing from spring 2000 to spring 2008. The BV were available for the three production traits: milk volume, fat and protein yields. In this dataset, bulls born between 1997 and 2003 had at least one PA evaluation and a subsequent progeny test.

The BV were adjusted for genetic group solutions to yield the additive genetic component of the BV. Let BV_{PA} and BV_{PT} be the BV of bulls based on parent average and progeny test respectively, with corresponding reliabilities R_{PA} and R_{PT}. The variance of the difference $BV_{PT} - BV_{PA}$ is proportional to the change in reliability, $R_{PT} - R_{PA}$. To standardise the change in BV we define a z-score for bulls

$$Z_{bull} = \frac{BV_{PT} - BV_{PA}}{\sigma_g \sqrt{R_{PT} - R_{PA}}} \tag{1}$$

where σ_g is the genetic standard deviation. Given zero bias, we would expect about 95% of the z-scores to lie in the range ±2. Similarly we define a z-score for dams of bulls as

$$Z_{dam} = \frac{BV_{dam} - \overline{BV}_{dam}}{\sigma_g \sqrt{R_{dam}}}$$

---

[*] LIC, Private Bag 3016, Hamilton 3240, New Zealand

where $\overline{BV}_{dam}$ is the contemporary group average for bull dams. Bulls were classified according to herd of origin and whether or not they were born through ET. The herd of origin was the bull's birth herd if that herd had supplied at least 5 bulls for progeny testing over a period of 5 or more years otherwise bulls were assigned to a common herd.

The dependent variable was Z_{bull}, and the explanatory variables were Z_{dam} fitted as a covariate, ET as fixed effect and herd as a random effect. Fitting herd as random avoided overestimation of herd effects using least squares where bull numbers are small and facilitated estimation of the variation due to herd effects.

A pedigree index (PI) was also considered as a substitute for PA in equation (1). The PI is based on the sire stack down the maternal side of the pedigree and is therefore not affected by potential biases in the bull dam evaluations. The BV_{PI} was calculated as $0.5BV_{sire} + \sum_{i\geq 2} 0.5^i BV_i$ where BV_i is the estimated BV of the male ancestor i generations back through the maternal side of the pedigree where, for example, $i = 2$ refers to maternal grandsire. The corresponding reliability is $R_{PI} = 0.25R_{sire} + \sum_{i\geq 2} 0.25^i R_i$

Results and discussion

The average values for Z_{bull} are shown in table 1. The columns of table 1 represent the genetic evaluations at the end of the season of the bull's birth year through to the evaluation as a 4-yr-old prior to the first lactation of the bull's daughters. The bias tends to decrease with time and the z-score averages -0.43 to -0.81 units. Assuming $R_{PA} = 0.35$ and $R_{PT} = 0.85$ this corresponds to PA overestimation of 3 to 4 kg in protein BV. The trend in PA bias over years could reflect general drift in evaluations (Powell et al. (2004)) which may include some parentage error associated with sires of sons as they accumulate daughter numbers.

Table 1: Mean values of Z_{bull}

	bull age (yr)			
	1	2	3	4
no. bulls	1963	2335	2685	3010
milk	-0.78	-0.71	-0.64	-0.53
protein	-0.81	-0.73	-0.67	-0.55
fat	-0.59	-0.52	-0.51	-0.43

The values corresponding to those in table 1 when using PI as a substitute for PA were generally positive suggesting some underestimation in this case. This is presumably due to the missing positive Mendelian sampling component for bull dams. The z-score values averaged 0.03, 0.26 and 0.33 for milk, protein and fat respectively when using PI.

About 29% of the bulls resulted from embryo transfer. The magnitude of the bias was greater for ET bulls: the z-score for ET bulls compared with non-ET bulls averaged -0.11, -0.14 and -0.15 for milk, protein and fat respectively. For example, the protein BV of young bulls born to ET is overestimated by about 0.9kg. The regression of Z_{bull} on Z_{dam} was fitted including separate intercepts for ET. The regression estimates, averaged over bull age, are presented in table 2. These effects were generally significant sources of variation. The ET effect is somewhat reduced after fitting the dam covariate. The covariate adjustment can be used to moderate the bias resulting from extreme dams.

Table 2: Regression estimates for Z_{bull}

trait	intercept	Z_{dam}	ET
milk	-0.59	-0.19	-0.02
protein	-0.56	-0.09	-0.11
fat	-0.36	-0.25	-0.06

Based on the REML likelihood ratio test, the random effect of herd made a significant contribution to the model. The variance of the herd effect for the z-score of protein yield averaged 0.05. This variation could be partly generated by a management group effect for bull dams that is not taken into account in the national evaluation scheme. The standard deviation of the herd effect is about 1.4 kg on the protein BV scale. Cassandro et al. (2002) also found significant herd of origin effects on bull proofs.

For a young bull with genomic information, assume that the reliabilities of the parent average and the genomic BV are 0.35 and 0.6 respectively. Then the reliability of the genomic information, marginal to the parent average is $R_G = \frac{0.6-0.35}{1+0.6*0.35-2*0.35} = 0.49$ (Harris and Johnson (1998)). Using selection index, the weighting applied to the parent average in the genomic BV would be approximately $\frac{1-R_G}{1-R_G R_{PA}} = 0.6$. Thus the adjustment for PA bias in a young bull with a genomic BV would be approximately 60% of the values discussed above. This procedure can be applied progressively down the pedigree. For example, for a young bull whose sire is genomically selected but without progeny, the BV of the sire will in turn be influenced by PA bias and so the correction scaling factor will be of the order of $0.6 * (1 + 0.6) = 0.96$.

Conclusion

These results quantify the extent of overestimation of BV of young bulls based on parent average in the New Zealand Animal Evaluation system. The effects considered suggest some overestimation may be due to management of bull dams which is not accounted for by contemporary grouping in the national genetic evaluation. Selection of bull dams from the more variable herds should not be a problem as heterogeneity of variance is adjusted for in the national evaluation. The PI as a substitute for PA suggests that most of the bias is derived from bull dam records. The cow to bull pathway is the most intensely selected pathway which also raises the possibility of non-linearity of the regression of offspring on parent or epistatic gene interactions (Bulmer (1971); Robertson (1976); Smith (1984)).

The current proposal for New Zealand Animal Evaluation is to adjust the PA evaluation of all young bulls according to the regression estimates in table 2. It is not clear at this stage whether an adjustment for herd can be implemented in a production system. In the future, young bulls screened on the basis of a SNP panel will likely originate from a wider range of herds. An alternative proposition that requires further consideration is the use of the pedigree index as a substitute for the parent average.

References

Bulmer, M. G. (1971). *Am. Nat.*, 105:201–11.

Cassandro, M., Dal Zotto, R., Gallo, L., Carnier, P., and Bittante, G. (2002). *Ital. J. Anim. Sci.*, 1:265–274.

Harris, B. L. and Johnson, D. L. (1998). *J. Dairy Sci.*, 81:2723–2728.

Pedersen, G. A., Christensen, L. G., and Petersen, P. H. (1995). *Acta Agric. Scand.*, 45:26–31.

Powell, R. L., Sanders, A. H., and Norman, H. (2004). *J. Dairy Sci.*, 87:2614–2620.

Rensing, S., Pasman, E., and Reinhardt, F. (2009). In *Proc. Interbull Meeting, Barcelona, Spain*, pages 123–126.

Robertson, A. (1976). In *Proc. Int. Conf. Quant. Genet. E. Pollak, O. Kempthorne and T. B. Bailey, Jr., ed. , Iowa State Univ. Press, Ames.*, pages 297–304.

Smith, C. (1984). *Res. Dev. Agric.*, 1:79–85.

van der Werf, J., Meuwissen, T., and De Jong, G. (1994). *J. Dairy Sci.*, 77:3174–3184.

Zwald, N. R. and Weigel, K. A. (2002). *J. Dairy Sci.*, 85, Suppl. 1:32.