

# ROBUSTNESS OF SELECTION RESPONSE PREDICTED FROM BLUP BREEDING VALUES

G.M. Macbeth

Queensland Department of Primary Industries  
Animal Research Institute  
665 Fairfield Road, Yeerongpilly, Qld 4105  
Australia.

## SUMMARY

The robustness of predicting selection response from BLUP using  $R = ir_{\hat{A}A}\sigma_A$  was measured as the difference in response determined from a simulation model when variations in accuracy occur. Selection response was overestimated by 25%, 11% and 2% when average accuracies ( $\bar{r}$ ) were 0.3, 0.4 and 0.5 with accuracy variation sampled using a variance of 0.125<sup>2</sup>. When accuracies were sampled over a uniform distribution with ranges 0 to 0.6, 0 to 0.8 and 0 to 1 overestimates were 64%, 25% and 4%.

## INTRODUCTION

A method of selection which has gained recent popularity is based on Best Linear Unbiased Prediction (BLUP) which efficiently utilises performance information from relatives. The accuracy of each animal's estimated breeding value (*ebv*) will therefore differ with the amount of pedigree information available. As the variance of accuracies increase so does the range of standard errors of prediction of *ebv*'s. It is from this range that outliers will be preferentially selected, having a large positive deviation of their *ebv*'s from the expectation of their *ebv*'s (Macbeth 1994). In this case it is thought that the true response to selection predicted by  $R = ir_{\hat{A}A}\sigma_A$  (Nicholas 1987) would be overestimated and that the deterministic equation would not be suitable for estimating BLUP selection response.

A simulation model was developed to test the theory that under situations where accuracies vary the deterministic equation overestimates true selection response.

## METHOD

The usual formula for estimating selection response is given by equation 1 where  $r_{\hat{A}A}$  is the correlation between the selection criteria ( $\hat{A}$ ) and true breeding value ( $A$ ). When varying amounts of pedigree information are used in BLUP procedures the selection criteria are *ebv*'s with variations in accuracies ( $r_{\hat{A}A}$ ) among them.

$$R = ir_{\hat{A}A}\sigma_A \quad \dots (1)$$

The response predicted by equation 1 is compared to that from a population with simulated variations in accuracy. A breeding population of 10,000 was simulated with animal  $j$  having true breeding values ( $A_j$ ), estimated breeding values ( $\hat{A}_j$ ), and accuracies ( $r_j$ ) derived using the methods of Klieve (1993) and Macbeth (1994) using:

$$\hat{A}_j = r_{\hat{A}A}^2 A_j + N(0, 1)\sigma_j \quad \dots (2)$$

where  $\sigma_i = [r_i^2 \sigma_A^2 (1-r_i^2)]^{0.5}$  ... (3)

$A_j \sim N(0, \sigma_A^2)$  ... (4)

and  $r_j \sim N(\bar{r}, \sigma_r^2)$ . ... (5)

The best 5% of  $\bar{A}$ 's from 50 simulation runs were selected and the mean of their related  $A$  values ( $\bar{A}_s$ ) and  $r$  values ( $\bar{r}_s$ ) determined where  $\bar{A}_s$  measures the response from the first generation of selection.

Given  $\sigma_A = 1.0$ ,  $i = 2.06$  (from a selected proportion  $p = 5\%$ ) the robustness of equation 1 was measured as the error of predicting selection response from that determined by simulation as:

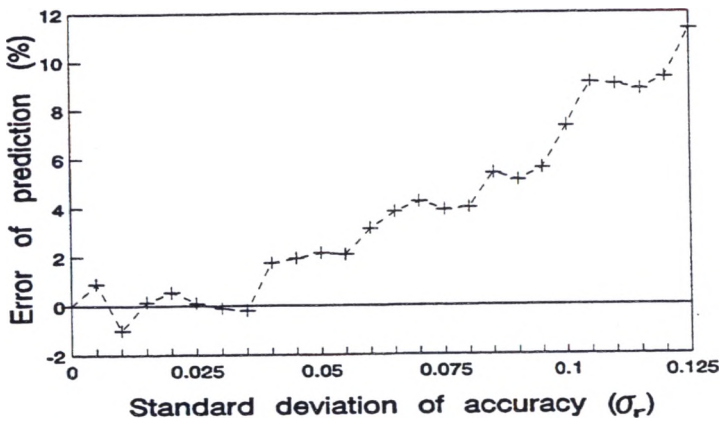
$$E = 2.06 \bar{r}_{\bar{A}\bar{A}} \bar{A}_s - 1 \quad \dots (6)$$

Additional simulations were run when  $r_j$  records were sampled from a uniform distribution. Since accuracies were sampled from a symmetrical distribution, mean accuracies of selection candidates  $\bar{r}$  equals the correlation  $r_{\bar{A}\bar{A}}$ .

### RESULTS

When the selected proportion ( $p$ ) is 5% and  $\bar{r} = 0.4$  the error in predicting response from equation 1 ( $E$ ) seemed to increase exponentially as the standard deviation of accuracy measurements increased (Figure 1). When accuracy variance  $\sigma_r^2 = 0.125^2$  and  $p = 1\%$ , 10%, and 20% the error ( $E$ ) was 10%, 6% and 1%. Response from equation 1 equalled the simulated response when  $\sigma_r = 0$ .

Figure 1. Error of predicting selection response from equation 1 when the proportion selected  $p = 5\%$  and mean accuracy of  $ebv$ 's  $\bar{r} = 0.4$ .



The error of prediction ( $E$ ) increased as the mean accuracies of  $ebv$ 's decreased and as the variation of accuracies of  $ebv$ 's increased (Table 1). The average accuracies of selected candidates ( $\bar{r}_s$ ) tended to be larger than the population mean in which they were selected ( $\bar{r}$ ).

Table 1. Average accuracies ( $\bar{r}_s$ ) and true breeding values ( $\bar{A}_s$ ) of animals selected and the error of predicting selection response ( $E$ ). Accuracies were simulated having different means ( $\bar{r}$ ) and distribution patterns and when the proportion selected  $p = 5\%$ .

Accuracy variation	$\bar{r}$	$\bar{r}_s$	$\bar{A}_s$	$E$ (%)
0.050*	0.3	0.322 ± .001#	0.574 ± .007	8
	0.4	0.414 ± .001	0.810 ± .008	2
	0.5	0.509 ± .001	1.023 ± .010	0
0.075*	0.3	0.344 ± .001	0.547 ± .007	13
	0.4	0.430 ± .001	0.798 ± .008	3
	0.5	0.517 ± .001	1.031 ± .010	0
0.125*	0.3	0.399 ± .001	0.494 ± .007	25
	0.4	0.472 ± .001	0.740 ± .007	11
	0.5	0.543 ± .001	1.007 ± .007	2
0.0 to 0.6**	0.3	0.649 ± .001	0.377 ± .007	64
0.0 to 0.8**	0.4	0.644 ± .001	0.661 ± .007	25
0.0 to 1.0**	0.5	0.627 ± .002	0.995 ± .008	4

# Standard error

\* Variation in accuracies sampled from normal distribution ( $\sigma$ , indicated)

\*\* Accuracies sampled from a uniform distribution (range indicated)

## DISCUSSION

Although there is no quantitative theory to determine the relative differences between selection on individual phenotype and BLUP estimated breeding values, stochastic simulation methods have shown an advantage of BLUP selection (Quinton et al 1992, Belovsky and Kennedy 1988). An easy method of estimating selection response, without having to resort to computer simulation, would be to use a deterministic approach using equation 1. This could be useful in feasibility studies of potentially new breeding schemes. However as shown this can be misleading when breeding values of selection candidates vary.

There are a number of different causes of the variation in accuracies of estimated breeding values, the most obvious being differences in pedigree records which can result in large variations of accuracy (Satoh et al 1992). Accuracies of sires may be as high as 0.99 due to extensive progeny records with medium heritable traits such as 0.25 for milk production (A.D.H.I.S. 1993). Different numbers of records on the same animal could also have a large influence on the variation of accuracies. As calculated in Avalos and Smith (1987) the accuracy of litter size increased from 0.36 with two litter records to 0.63 with four litter records. Across herd selection with different record types could also contribute to variations in accuracy. With these and possibly other factors affecting the variation in accuracy the frequency distribution of accuracies can be far from normal and may appear to have a much flatter and more evenly distributed pattern as was found in 1993 A.D.H.I.S. milk production data. Without prior knowledge of the frequency distribution of accuracies, the error in predicting BLUP selection response from equation 1 therefore cannot be obtained. This was evident due to the large difference in the error of predicting response between normal and uniform frequency distributions of accuracies. However it does appear that, when average accuracies are high (eg  $\bar{r} > 0.5$ ), the response predicted by equation 1 is in error by less than 4% and that this may be suitable for estimating approximate response.

The average accuracies of candidates selected tended to be greater than that of the population average. This was more pronounced as the population accuracies declined from 0.5 to 0.3. A possible explanation for this is that the standard error of prediction of estimated breeding values ( $\sigma_{ebv}$ ) is largest when  $r=0.7$  (determined from differentiating  $\sigma_{ebv}$  with respect to  $r$ ). Outliers on  $ebv$  selection therefore tended to be chosen from animals having accuracies in the vicinity of  $r=0.7$ . Groups of animals selected having higher accuracies are more likely to be chosen from animals with higher inbreeding levels. This would result in a reduction of additive genetic variance and a reduction in asymptotic selection response (Wray and Hill 1989).

Using the same simulation methods as this study it was shown that, in populations where variation in low accuracies occur, outliers could be avoided by using indexed breeding values (Macbeth 1994). The index on each animal  $j$  being determined from  $(\hat{A}_j - b \cdot \sigma_j)$  where  $b$  is a constant chosen to optimise selection response under different population and selection conditions. This improved the rate of selection response and at the same time selected animals with lower accuracies than the population average.

#### REFERENCES

- A.D.H.I.S. (1993) ABV book. Australian Dairy Herd Improvement Service. May 1993. Glen Iris. Victoria.
- AVALOS, E. and SMITH, C. (1987) *Animal Production*, 44 : 153-164.
- BELONSKY, G.M. and KENNEDY, B.W. (1988) *J. Anim. Sci.* 66 : 1124-1131.
- KLIEVE, H.M., KINGHORN, B.P. and BARWICK, S.A. (1993) *J. Anim. Breed. Genet.* 110 : 1-12.
- MACBETH, G.M. (1994) *J. Anim. Breed. Genet.* (In Press).
- NICHOLAS, F.W. (1987) *Veterinary Genetics*. Oxford University Press. New York.
- QUINTON M., SMITH, C. and GODDARD, M.E. (1992) *J. Anim. Sci.* 70 : 1060-1067.
- SATOH, M., NISHIDA, A. and FURUKAWA, T. (1992) *Animal Science and Technology*. 63 : 457-461.
- WRAY, N.R. and HILL, W.G. (1989) *Animal Production*, 49 : 217-227.