

ACCURACY OF ESTIMATED BREEDING VALUES

B. TIER, M. SCHNEEBERGER and K. HAMMOND

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

INTRODUCTION

Information from observations on an animal and its relatives is used to predict Breeding Values for that animal. An animal's Breeding Value (BV) is defined as the mean performance of many offspring. During meiosis chromosomes duplicate and then separate to form gametes. The average genetic worth of the gametes is half the animal's BV. Gametes from two parents combine to form a new individual - hence the expected BV of an offspring is half the sum of its parents' BVs. For non-inbred animals the variance of the genetic values of the gametes is one quarter of the genetic variance. For inbred animals it is less and depends on the degree of inbreeding. All gametes from a completely inbred organism are identical. This is a feature exploited by plant breeders but completely inbred animals are unachievable because of fitness constraints.

In an analysis using Best Linear Unbiased Prediction (BLUP, Henderson 1973) Estimated BVs (EBVs) are computed simultaneously and in relation to each other. The EBVs rank the animals and predict the differences between them. Some of the major features of a BLUP analysis are to separate the genetic effects from systematic environmental effects by contrasting an animal's performance with that of its contemporaries (i.e. animals of a similar age treated in a like manner), and to simultaneously adjust genetic effects for unequal competition (e.g. resulting from mating best to best and from sequential culling).

With few observations on an animal and its relatives, its EBVs may poorly reflect the mean of its gametes. As more data are recorded then the degree of certainty associated with the EBVs improves. The contribution that each observation makes toward an animal's EBV depends upon: 1) The animal's relationship with the observed animal, which is defined by the pedigree and incorporated into the BLUP analysis via the inverse of the numerator relationship matrix. 2) The genetic correlation between the trait which was observed and the trait for which an EBV is to be predicted - defined by the genetic (co)variance matrix used in the analysis. This includes the heritability of the measurement itself. 3) The value of other observations which contribute to the EBV. These may be observations on the animal itself, its relatives or its contemporaries or their relatives. 4) The number of contemporaries that an animal is compared to and its relationship to them.

The Accuracy (ACC) of an EBV, expressed as a percentage, is a measure of the confidence that can be placed in the prediction of the animal's BV for a particular trait. It can be related to the prediction error variance (PEV) of each EBV. The percentage accuracy for each EBV is given by the formula

* AGBU is a joint venture between the University of New England and the NSW Department of Agriculture and Fisheries.

$$ACC=100(1-PEV/\sigma_g^2)^{1/2},$$

where σ_g^2 is the genetic variance for the trait. The PEVs are never greater than the genetic variance for the trait and hence accuracies are never greater than 100 per cent. However, to compute the PEVs requires inverting a large matrix - the coefficient matrix of the Mixed Model Equations utilised in the BLUP procedure. For most applications obtaining this inverse is a time-consuming and very expensive procedure.

There are some underlying assumptions about BLUP analyses that need to be considered in relation to the accuracies. They are: 1) Data are measured consistently - i.e. are free of bias. 2) The genetic (co)variances are the *true* parameters. 3) There are no errors in pedigree records. 4) The model employed is correct - i.e. all effects are fitted properly. At least some of these assumptions will not be met in practise. In all cases estimates of the genetic parameters are employed since the true parameters are unknown. This may affect accuracy more than it affects the EBVs. If either assumption 1 or 4 is not met then it is likely that the accuracies will be overestimated. If assumption 3 is incorrect then the accuracies on some animals will be overestimated and on others underestimated.

Because the calculation of accuracies is an expensive and time-consuming process they are generally not computed. However there are many methods for computing approximate accuracies (Van Raden and Freeman, 1985). These range from simple methods which recognise the numbers of offspring and observations that an animal has through inverting the diagonal of the coefficient matrix to computing the inverse of a modified coefficient matrix. Most of these methods have not been applied to multiple-trait animal models, the BLUP approach used in BREEDPLAN (Nicol et al. 1985). Most of them also overestimate the accuracies for EBVs but do so inconsistently - begging the question: "how accurate are these approximate accuracies?". The answer to this question is beyond the scope of this paper.

VALUE OF DIRECT INFORMATION

We examine how information observed on a trait influences the accuracy of EBVs for that trait. In the examples the accuracies of EBVs for animals and their sires and dams were computed for a range of contemporary group sizes (2, 5, 10, 25 and 50 members), a range of heritabilities (0.1, 0.25 and 0.5) and a range of previous levels of accuracy for *all* the parents in the contemporary group (0, 40, 70 and 90%).

Contemporary groups were designed so that each animal had a unique dam. In the small contemporary groups (2 and 5 animals) each animal also had a unique sire, but in the large groups (10, 25 and 50 animal) five sires were used in equal proportion, e.g. in a contemporary group of 25 animals there were five half-sib cohorts of 5 animals.

The accuracy of the parental EBVs prior to the addition of each contemporary group was set to one of four levels - 0 (Z), 40 (L), 70 (M) or 90% (H). These levels of accuracy identify the contribution from observations on the parents and other relatives. The accuracy of the EBVs of progeny from parents with these levels of accuracy without any further information (no measurements on the progeny themselves) are 0, 28, 50 and 64 percent respectively, according to the formula

$$ACC = 100\{1 - [(0.25 PEV_S + 0.25 PEV_D + 0.5\sigma_g^2)/\sigma_g^2]\}$$

where PEV_S , PEV_D are the parental PEVs.

The results from this analysis are presented in Figure 1. Accuracy of the EBVs of the observed animals increases with the size of the contemporary groups, heritability and previous level of accuracy of their parents. At moderate to high levels of heritability contemporary groups of five animals provide almost as much information on the observed animals as the larger groups. (Note, this assumes all sires are equally

represented in each contemporary group). The higher accuracies of the sires' EBVs compared with those of the dams in the larger contemporary groups reflect the value of their additional progeny. The accuracies of the parents' EBVs indicate how much additional information is required to increase the level of accuracy when the previous level is high compared to when it is low. This illustrates how the value of any observation is affected by contributions from other relatives.

VALUE OF CORRELATED INFORMATION

This section demonstrates how information observed on one trait influences the accuracy of the EBV for a correlated trait. The same contemporary group structures and levels of accuracy of parental EBVs described in the previous section were used in this analysis. Two levels of heritability (0.15 and 0.40) of the observed trait and three genetic correlations (0.25, 0.5 and 0.75) were used.

The accuracies of the correlated EBVs for animals on which the additional information was observed are presented in Figure 2. These accuracies increase as contemporary group size, heritability of the observed trait and genetic correlation between the observed and predicted traits increase, as these variables influence the genetic information content of a data set. Only when the genetic correlation is high or there is little other information on the animal's relatives does the observation on one trait significantly improve the level of accuracy of the EBV of a correlated trait. Of course, in the multiple-trait BREEDPLAN system the information from all correlated traits is utilised simultaneously.

ACCURACIES IN PRACTICE

In this study the effect of additional information from increased parental information, contemporary group size, heritability and genetic correlation on accuracy was illustrated holding everything else constant. Sires and dams were assumed to be unrelated and equally represented in all contemporary groups. The effects of common environment artificially augmenting the true genetic relationship between relatives were neglected and the amount of information from relatives in the pedigree was the same for all animals. Field data sets, however, are generally highly unbalanced; hence, the effect of additional information on accuracy will be more complex in practice.

RISK AND ACCURACY

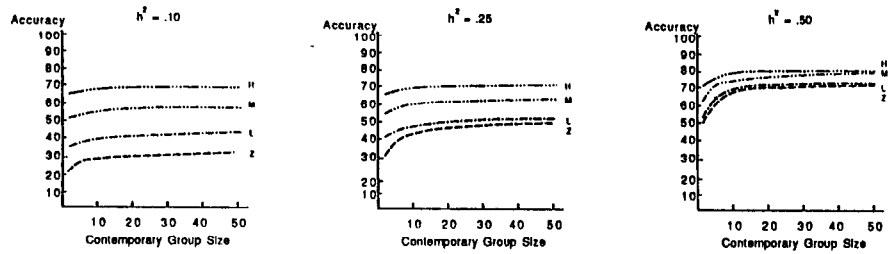
Accuracies of EBVs provide a measure of risk with high accuracy representing low risk. Portfolio theory can be applied to make selection decisions according to a decision maker's attitude towards risk, i.e. his utility function. This implies making a decision between "portfolios" of animals with high expected values (mean EBVs) and high variances and portfolios with lower mean EBVs and smaller variances. Smaller portfolio variances can be achieved by selecting animals which have more accurate EBVs and/or by selecting a diverse group of animals (Schneeberger et al. 1982).

However, Accuracies and EBVs must be kept in their relative perspective. For example, genetic change is achieved primarily from use of the EBV's themselves. Genetic change does not occur in Accuracy! The EBVs are already regressed according to the amount of information involved in their calculation.

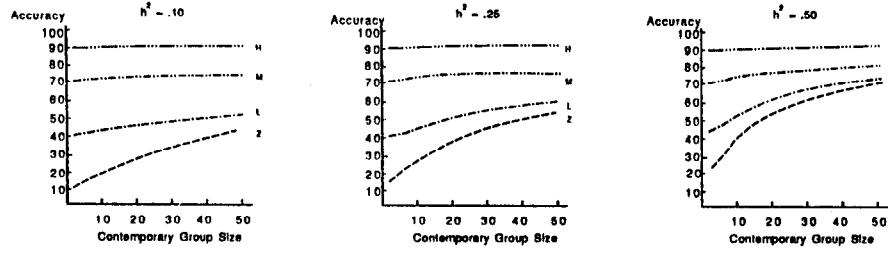
REFERENCES

- HENDERSON, C.R. (1973) Proc. Anim. Breed. Genet. Symp. in Honor of Dr. J.L. Lush: 10. Amer. Soc. Animal Sci. and Amer. Dairy Sci. Assoc., Champaign, Illinois.
- NICOL, D.C., GRASER, H.-U., TIER, B. and HAMMOND, K. (1985) Proc. AAABG 5th Conf. Sydney: 151.
- SCHNEEBERGER, M., FREEMAN, A.E. and BOEHLJE, M.D. (1982) J. Dairy Sci. 65:404.
- VAN RADEN, P.M. and FREEMAN, A.E. (1985) J. Dairy Sci. 68:2123.

I) Animal



II) Sire



III) Dam

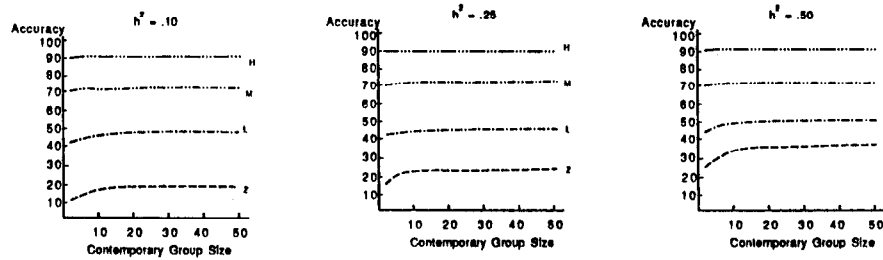
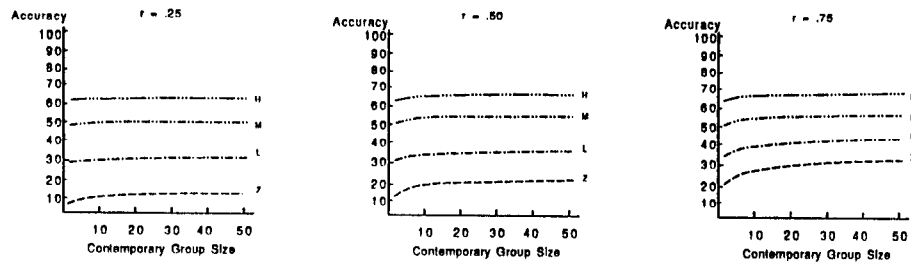


Figure 1: The accuracy of an EBV for an animal and its sire and dam resulting from the direct observation of the trait. Accuracies are illustrated for observations on the animal with different numbers of contemporaries, for traits with different heritabilities (h^2) and for different previous levels of parental accuracy (reflecting the contributions from observations of other relatives, Z=0, L=40, M=70 and H=90%).

I) $h^2 = .15$



II) $h^2 = .40$

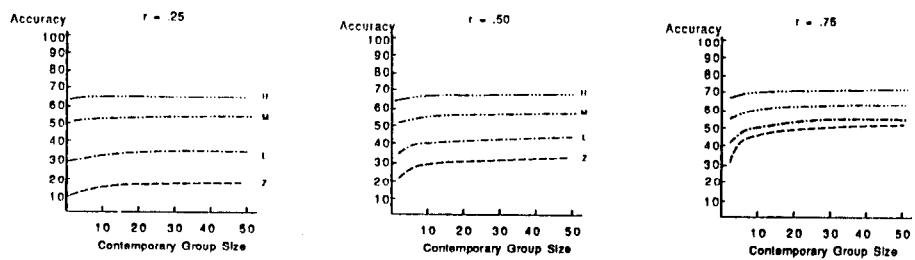


Figure 2: The accuracy of an EBV for an animal resulting from the observation of a correlated trait. Accuracies are illustrated for observations on the animal with different numbers of contemporaries, for traits with different heritabilities (h^2), different genetic correlations (r) and for different previous levels of parental accuracy (reflecting the contributions from observations of other relatives, Z=0, L=40, M=70 and H=90%).