MEASURING CONNECTEDNESS: CONCEPT AND APPLICATION TO A LARGE INDUSTRY BREEDING PROGRAM

P. K. Mathur, B. P. Sullivan and J. P. Chesnais

Canadian Centre for Swine Improvement
Central Experimental Farm, Building 54, Maple Drive, Ottawa, ON, K1A 0C6, Canada

INTRODUCTION

When selection is between animals raised in different herds, the accuracy of comparisons between genetic evaluations can be influenced by the degree of connectedness among these herds. Evaluations from herds that are poorly connected may be severely underestimated or overestimated. Several methods have been proposed to evaluate connectedness. However, connectedness can be defined in different ways. In addition, some methods are very demanding computationally, which makes their routine application difficult. The objective of this paper is to briefly review the concept of connectedness and evaluate the performance of a connectedness rating method that has been used in the Canadian swine improvement program for the last four years.

CONCEPT AND ESTIMATION OF CONNECTEDNESS

Connectedness, because it measures the accuracy of comparisons across EBVs, is a statistical measurement rather than one of genetic relationship. Unrelated groups of animals are "connected" if they are tested in the same management group. Two management groups can be connected through a third group even if there are no direct genetic links between them. Central test stations can increase connectedness without increasing relationships among herds. Methods that are based on the degree of genetic relationships between animals rather than on a statistical measurement of connectedness i.e. those based on "gene flow" or on additive genetic relationships will often give inferior results, as shown by Kennedy and Trus (1993).

Since the main objective of measuring connectedness is to obtain an indication of the accuracy of comparisons between EBVs in different herds, a logical statistical measurement of the connectedness between two herds would be the average PEV of all pair wise EBV differences between the animals in the two herds. This is the standard put forward by Kennedy and Trus (1993) against which to evaluate all other methods. Other standards have been proposed, notably by Laloe et al (1996) who compared PEV with two other statistical measurements of connectedness: the squared correlation between prediction and true differences of genetic values (CD) and a connectedness index (IC) suggested by Foulley et al (1992) equal to the relative decrease in PEV when fixed effects are known. The authors conclude that CD combines aspects of genetic variability and PEV. However, if the primary objective of a connectedness rating is to identify herds where EBVs are poorly estimated in comparison to those of other herds so that remedial action can be taken, then a method which assesses only the accuracy of such comparisons would be most appropriate. PEV was therefore chosen as the basis for estimating connectedness.

Computing the average PEV of all pair wise EBV differences between animals for every possible pair of herds would be extremely time consuming. Therefore, an easier method was required. Kennedy and Trus (1993) confirmed through simulation that the variance of difference between herd effect estimates is very highly correlated (0.995) with the average PEV of differences between EBVs. Therefore, this variance can be used as a substitute measurement for PEV. Since a direct inverse of the entire set of mixed model equations would be difficult to compute, elements of the inverse were obtained one herd at a time.
Since $W W^{-1} = I$, each column of W^{-1} can be computed by solving $W w_i^{-1} = I_i$
where W is the coefficient matrix of mixed model equations, w_i^{-1} is a vector of W^{-1} for herd i and I_i is a vector with 1 for the herd effect and zero otherwise.

For the purpose of computing connectedness ratings, the vector of inverse elements was obtained for management group effects corresponding to the last six-month period in each herd. This vector contained the coefficients for the variances and covariances of estimates for management group effects. The variance of the difference between each pair of herd effects could be used as a measure of connectedness. However, the variance depends upon the size and structure of each pair of herds as well as the nature of the connections between them. An error variance of 0.8 mm^2 for backfat, for example, may correspond to either two large herds that are poorly connected, or two small herds that are well connected. To separate the notion of connectedness from the effects of herd size and structure, the connectedness rating between two herds was defined as the correlation between the estimates of herd effects, i.e.

$$CR = \frac{\mathrm{cov}(\hat{h}_1, \hat{h}_2)}{\sqrt{\mathrm{var}(\hat{h}_1)\,\mathrm{var}(\hat{h}_2)}}$$

In this manner, any reduction in accuracy associated with insufficient connectedness can be more effectively separated from that associated with insufficient herd or management group size. However, both connectedness as defined here and herd size have an important effect on accuracy of across-herd comparisons, as would be reflected in the variance of herd estimates or the PEV. If herd size is small, increasing management group size will do more for accuracy than increasing connectedness.

APPLICATION TO THE CANADIAN SWINE IMPROVEMENT PROGRAM

Data The above method was applied to data from the Duroc, Yorkshire, and Landrace breeds for backfat and age, and from the Yorkshire and Landrace breeds for litter size. Connectedness ratings (CR) were computed for each herd with every other herd on the national program for each breed and trait. CR between management groups in the last six months was used as a measure of connectedness between herds. CR were computed and reported officially every six months starting in 1997. The structure of the data used in the June 2001 evaluations is shown in Table 1.

Table 1. Data structure

Breed	Herds	Management Groups	Pigs	Equations
Duroc	67	7,512	299,854	377,443
Landrace	87	11,066	690,722	870,390
Yorkshire	101	13,684	907,741	1,152,855

An average CR was calculated for each herd as the average of a herd's CR with all other herds in the program. The average rating gives an indication of the accuracy of comparing EBVs from one herd to all others.

Connectedness Rating and accuracy of comparing EBVs across herds The relationship between the CR and the variance of differences between herd effect estimates was studied for lean growth (backfat) and for litter size in all breeds. Estimates for sow productivity in the Landrace breed are shown in Figure 1 as an example.

As expected, the variance decreases as the CR increases. The relationship is non-linear. The variance is high when the CR is low and decreases gradually with the increase in CR. These results suggest that the CR can be used to predict the accuracy of comparisons between EBVs of specific pairs of herds and the non-linear relationship can be used to determine the optimum level of CR required for such comparisons.

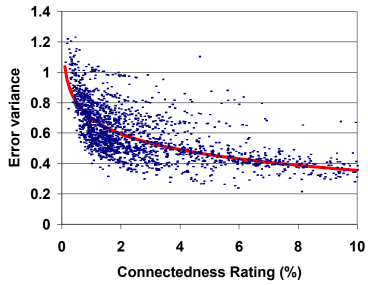

Figure 1. CR versus variance of difference between herd effect estimates

Effect of sample size The study of the relationship between CR and herd size or variance has shown that, as expected, CR is less strongly related to these factors than the variance of differences between herd effects. However, when herd size is small (less than 50), CR is lower and still tends to decrease with herd size. While this is desirable in practice, research is required to determine whether this reflects a lower level of exchanges among smaller herds or some residual dependency between CR and herd size.

Trends in connectedness for herds on the Canadian Swine Improvement Program Averages in CR between herds from 1997 to 2001 are shown in Figure 2. Connectedness has increased substantially since connectedness evaluations were introduced. The increase is consistent with increased usage of common boars through artificial insemination, as program participants with low CR have followed this recommendation.

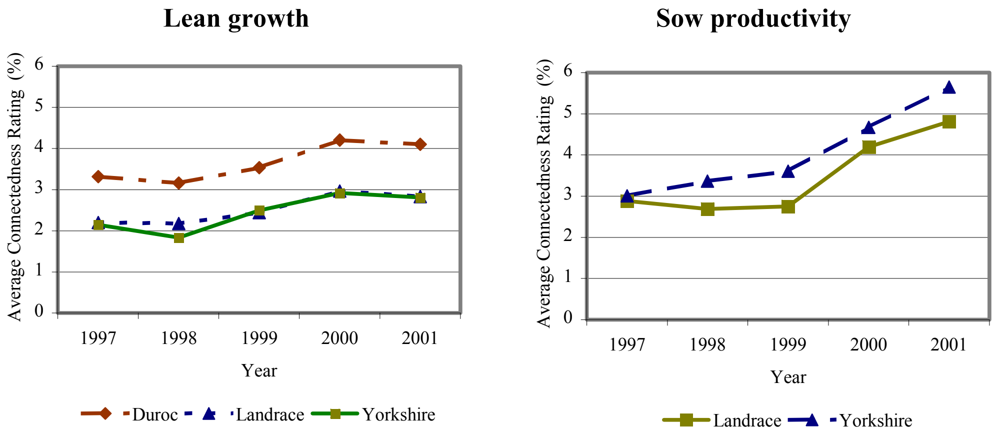

Figure 2. Trends in connectedness for major swine breeds in Canada

CONCLUSIONS

CR appears to be a useful estimate of connectedness. As the correlation between estimates of recent herd effects, it is less strongly linked to herd size and variation than the PEV of pair-wise comparisons of EBV across herds, and therefore reflects more closely the concept of connectedness per se. It also matches the expectations of breeding businesses about connectedness based on their knowledge of the level of genetic exchanges among herds in the program, and as such is well accepted by the industry. The method can be easily applied to other livestock species.

REFERENCES

Foulley, J.L., Hanocq E.,Boichard, D (1992) *Genet. Sel. Evol.* **24** : 315-330
Kennedy, B. W. and D. Trus (1993) *J. Anim. Sci.* **71** : 2341-2352.
Laloe, F., Phocas, F. and Menissier, F (1996) *Genet. Sel. Evol.* **24** : 315-33