

# **Estimation of the degree of Connectedness between herds or management groups in the Canadian swine population**

P. K. Mathur, B. Sullivan and J. Chesnais  
Canadian Centre for Swine Improvement, Ottawa, Canada

## **1. Introduction**

To practice effective selection, we must be able to compare the breeding values (EBVs) of animals in different contemporary groups, i.e. animals probed at different times within the same herd, or in different herds. The EBVs of animals from different contemporary groups or herds are comparable due to the method used in producing the evaluations (BLUP). However, the accuracy of these comparisons depends upon the degree of connectedness among the groups or herds.

The accuracy or “repeatability” associated with each EBV reflects the amount of information on which the EBV is based. It gives a general estimate of the reliability of the EBV as a measure of true genetic value of the animal, but it is not necessarily a good indicator of the accuracy of comparisons with EBVs in other herds. If there are few or no genetic exchanges between a herd and the remainder of the tested population, the EBVs of all animals in that herd cannot be compared accurately to the EBVs estimated in other herds even when they have a high repeatability. When selecting animals across herds or management groups within a herd, decisions must be made on whether EBVs can be compared without excessive risk. These decisions must be based on the degree of the connectedness between the groups or herds. The lower the connectedness, the higher the risk. Therefore, it is important to measure the degree of connectedness among herds, and if necessary bring it to a level that allows comparison of EBVs with reasonable accuracy.

There is no well established procedure for measuring the degree of connectedness. Therefore, research was carried out, first through a contract with Dr. L.R.Schaeffer from the University of Guelph, and later at CCSI. The objectives of the research were: 1) to develop an accurate and practical method of measuring the degree of connectedness among herds or contemporary groups 2) to apply the method to all herds in the national genetic evaluation program and 3) to develop recommendations to ensure a minimal level of connectedness among the herds on the program.

## **2. The concept of connectedness**

Connectedness, because it measures the accuracy of comparisons among EBVs, is a statistical measurement rather than one of genetic relationship. Unrelated animals are connected if they are tested in the same management group. This is also true of unrelated animals in different herds. Take the example of 3 herds (A, B and C). Sire 1 has progeny in herds A and B, and sire 2 in herds B and C. If sire 1 and 2 have progeny in a common management group in herd B, they are “connected”, i.e. their genetic values can be compared with some degree of accuracy. Because sires 1 and 2 have progeny in herds A and C, these two herds are also

connected. This is true even if they do not contain animals that are genetically related. Therefore, the term “genetic connectedness” is inappropriate. Genetic relationships between animals in different herds increase the degree of connectedness, but are not necessary for connectedness to exist. There are other ways to create connectedness. For example, the degree of connectedness between two herds will increase whenever other herds import genetic material from them and compare it in the same environmental conditions. Central test stations used to provide a powerful way to increase connectedness although they did not markedly increase the degree of relationships among herds (since few station tested boars were sold to other nucleus herds). Methods that are based on the degree of genetic relationships between animals rather than on a statistical measurement of connectedness (e.g. the “gene flow” method in Kennedy and Trus, 1993 or estimates based on additive genetic relationship in Banos and Cady, 1988) will often give inferior results.

Since the 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 pairwise EBV differences between the animals in the two herds. However, computing this statistic for all herds in the program would be extremely time-consuming, so that an easier method is required.

The EBV of an animal with its own record depends, among other factors, upon the deviation between its performance and the estimate of the effect for the herd in which the animal was tested. Therefore, the PEV of the difference between the EBVs of two animals in different herds or groups can be separated into two components: one associated with the animals and their relatives (which depends upon the repeatabilities of each EBV, the degree of relationship between the two animals and covariances between animal and herd effects) and one associated with the herd effects themselves (accuracies of estimation of each herd effect, and the covariance between them). When calculating the average PEV of all pairwise differences between the EBVs of animals in the two herds, the component associated with the herd effects is the same for each comparison, while that associated with the animals and their relatives is different for each comparison and tends to cancel out over all comparisons. As a result, the average PEV over all comparisons tends towards the PEV of the difference between estimates of the two herd effects. Therefore, the PEV of the difference between the herd effects can be used as a measure of the degree of connectedness between two herds.

Kennedy and Trus (1993) confirmed through simulation that the PEV of the difference between herd effects is very highly correlated (0.995) with the average PEV of the differences between EBVs, and they proposed it as the method of choice to measure connectedness. However, they felt it would be difficult to apply to large populations with many management units because of the amount of computations involved. In addition to the computational challenge, the PEV of the difference between estimates of herd effects has another drawback. It depends on the size and structure of the two herds as well as on the nature of the “connections” between them. If the connectedness rating is the PEV itself, a rating of 0.8 mm<sup>2</sup> for backfat, for example, may correspond to either two large herds that are not well connected, or two smaller herds that are well connected. To separate the notion of connectedness from the effects of herd size and

structure, the connectedness rating was defined as the correlation between the estimates of the herd effects, i.e.

$$CR_{ij} = \frac{\text{cov}(\hat{h}_{ij})}{\sqrt{\text{var}(\hat{h}_i) \text{var}(\hat{h}_j)}} \quad (1)$$

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. The joint effect of these factors on the accuracy of the comparisons between EBVs from different herds can be formulated as follows:

$$\text{PEV}(\hat{h}_i - \hat{h}_j) = \text{var}(\hat{h}_i) + \text{var}(\hat{h}_j) - 2 \text{cov}(\hat{h}_{ij}) \quad (2)$$

$$= \text{var}(\hat{h}_i) + \text{var}(\hat{h}_j) - 2 CR_{ij} \sqrt{\text{var}(\hat{h}_i) \text{var}(\hat{h}_j)} \quad (3)$$

A rough approximation of the PEV based on management group size is then:

$$\text{PEV}(\hat{h}_i - \hat{h}_j) \cong \sigma_e^2 \left( \frac{1}{n_i} + \frac{1}{n_j} - 2 \frac{CR_{ij}}{\sqrt{n_i \cdot n_j}} \right) \quad (4)$$

where,  $\sigma_e^2$  = residual variance for the trait, and

$n_i$  and  $n_j$  are the number of records in the management groups or herds.

An approximation of the standard error of prediction (SEP) between herd estimates would be:

$$\text{SEP}(\hat{h}_i - \hat{h}_j) = \sigma_e \sqrt{\frac{1}{n_i} + \frac{1}{n_j} - \frac{2CR_{ij}}{\sqrt{n_i \cdot n_j}}} \quad (5)$$

Based on the above, the following steps were undertaken:

- examples were developed to investigate the behavior of the connectedness rating in various situations;
- research was carried out to find an efficient algorithm to compute the variances and covariances between estimates of herd effects for large data sets;
- the method was applied to all Duroc and Hampshire herds in Canada for backfat and age and all Yorkshire herds for litter size;
- the results were analyzed to determine what was a satisfactory level of connectedness rating for herds in the national program;
- recommendations were made on ways to increase the level of connectedness in herds where this was considered a problem;
- a measurement of connectedness for AI boars that could be used as part of the above recommendations was developed.

### 3. Application to simple examples

A comparison of the connectedness rating with the PEV of differences between breeding values is illustrated in appendix I for three simple examples. The examples are the same as examples 3, 4 and 5 in Kennedy and Trus except that the estimates of heritability genetic and residual variances for backfat are those in current use in the national swine improvement program.

These examples show that as the degree of relationship within herd increases, the PEV of within herd comparisons decreases but the PEV of across herd comparisons tends to increase. The PEV of across herd comparisons decreases when the connectedness rating increases. This decrease is due to a reduction in the variances of contemporary group effects and an increase in the covariances between them.

The effect of connectedness also depends upon the heritability of the trait. The heritability is different for age to 100 kg ( $h^2 = 0.32$ ) and for litter size ( $h^2 = 0.11$ ). This can make a marked difference in the comparison of EBVs. The effect in terms of the range of differences expected between EBVs is given in Table 1 for the above 3 cases. Here, the confidence range for the comparison of  $a_1$  and  $a_2$  within herd is 2.6 for backfat. If unrelated animals are compared within a herd ( $h_1$  or  $h_2$ ), in 95% of the cases the true difference in the means of the breeding values will be no more than  $\pm 2.6$  mm from the difference in their EBVs for backfat. This difference decreases to  $\pm 2.2$  mm if they are full sibs within the herd and to  $\pm 2.5$  mm if there are other related animals across herds. Similar trends are seen for age to 100 kg and litter size.

Table 1: 95% confidence range for differences between EBVs of animals in connected and disconnected herds

	Backfat (mm)			Age to 100 kg (days)			Litter size		
	CR = 0 %		CR= 25%	CR = 0 %		CR= 25%	CR = 0%		CR= 25%
	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3	Case 1	Case 2	Case 3
Within herds									
a1-a2	2.6	2.2	2.5	12	10	14	2.4	1.8	3.9
a3-a4	2.6	2.2	2.5	12	10	14	2.4	1.8	3.9
Between herds									
a1-a3	3.3	3.5	2.4	14	19	10	2.5	5.5	1.8
a2-a4	3.3	3.5	2.4	14	19	10	2.5	5.5	1.8
a1-a4	3.3	3.5	2.8	14	19	14	2.5	5.5	4.0
a2-a4	3.3	3.5	2.8	14	19	14	2.5	5.5	4.0

CR = Connectedness rating

#### **4. Computational method for large data sets**

One of the major problems in using this approach with large data sets is that a direct inverse of the entire set of mixed model equations is very difficult to compute. Hence, the following procedure was used, based on L. Schaeffer's suggestion.

$$w'w (w'w)^{-1} = I \quad \text{so,} \quad w'w (w'w)_i^{-1} = I_i \quad (6)$$

where,

$w'w$  = coefficient matrix of mixed model equations,

$I$  = identity matrix,

$I_i$  = a vector of the identity matrix corresponding to contemporary group  $i$  (a vector with 1 for the contemporary group and zeros),

$(w'w)_i$  = a vector of inverse elements for the  $i$ th contemporary group.

The vector  $(w'w)_i^{-1}$  was obtained for one contemporary group at a time through iteration (1000 rounds). These vectors were combined and the block of inverse elements corresponding to the most recent contemporary groups was extracted out. The inverse elements were the prediction error variances and covariances for the herd effects of interest.

The method can be used to obtain inverse elements for some rows or columns of any large matrix for which a direct inverse is not possible. It was first applied to small examples and exactly the same results were obtained with this method and with a direct inverse. More iterations were required for larger data sets as the number of iterations increased. Therefore, 1000 rounds of iterations were used. The computing time varied depending upon the number of equations extracted and the total number of equations (e.g. 12 hours and 20 minutes for 20 Hampshire herds, and 58 hours and 27 minutes for 71 Duroc herds). Nevertheless, the method makes it feasible to compute the inverse elements of relatively large matrices. It can be used to compute the exact prediction error variances and covariances for herds, litters, animals or any other effect included in the model, despite the very large total number of equations.

#### **5. Application to herds on the national program**

##### **5.1. Material**

The above method was applied to the national data base in the Duroc and Hampshire breeds for backfat and age, and to the Yorkshire breed for litter size. The connectedness ratings between herds were calculated using the last management group of each herd (pigs probed during the first half of 1997). Estimates of connectedness within herd were also obtained using the last two management group periods in order to study the relationship between connectedness and other variables for wide range of connectedness ratings. The number of animals and herds in each data set are shown in Table 2.

Table 2: Number of records analyzed

	Breed	Duroc	Hampshire	Yorkshire
	Trait	Backfat	Backfat	Litter size
Records		212,602	96,527	147,089
Animals		221,690	102,591	53,436
Equations		286,714	133,816	118,876
Contemporary groups		4,960	3,594	1,842
Herds with more than 10 records in the last contemporary group		71	20	77

## 5. 2. Connectedness ratings for pairs of herds

Connectedness ratings were calculated for each pair of herds in the program that had at least 10 records in the last contemporary group of the latest genetic evaluation. Examples of the results are given in the form of connectedness reports for four herds in Appendix II. These reports contain the connectedness of each herd in the program, with all other herds as well as the average connectedness of all herds.

Each of the four selected herds had high connectedness with few herds and low connectedness ratings with most other herds on the program. The ratings were usually higher with herds in the same region, but there were some good connections across regions as well.

These reports could be provided to each breeder for their own herds. At present the ratings are based on backfat records for the Duroc and Hampshire breeds. Similar reports will be available for the Yorkshire and Landrace breeds based on litter size records.

Maximum connectedness in the Hampshire breed (47.2 %) was observed between herd no. 684 and herd no. 1720. The structures of the contemporary groups for these two herds are shown in Table 3:

Table 3: Structure of Hampshire herds with maximum connectedness ratings

	Herd no. 684			Herd no. 1720		
	Total	common	% common	Total	common	% common
sires	10	6	60	13	6	46
litters	31	24	77	23	16	69
Pigs	60	45	75	40	25	62

There were 6 sires that had progeny in both herds. These 6 sires were used to produce 77% of the litters in herd 684 and 69% of the litters in herd 1720. Because of this, the two herds had strong genetic links and the highest ratings. Among other herds, the pairs of herds having connectedness rating below 3% did not have any sires in common during the last contemporary group period.

### 5.3. Average connectedness ratings

The average connectedness rating for a herd was defined as the average of its connectedness ratings with all other herds in the program. The values for Hampshire herds are given in Table 4 as an example and the average ratings for Duroc and Yorkshire herds are given in Appendix III and IV, respectively.

As can be seen by comparing individual herd reports (Appendix II) and these average, the averages connectedness rating of a herd can appear high because of very high connectedness ratings with a few herds. However, average ratings give an indication of the accuracy of comparing of EBVs from one herd to all others on the program. High connectedness with another herd allows effective selection of animals from that herd, while high average connectedness provides a broad basis for selection of superior animals from all other herds on the program. The measurement of connectedness is therefore very important for overall genetic improvement in a supernucleus structure such as the national swine improvement program.

Table 4: Average connectedness ratings for Hampshire herds

Herd name	Herd No.	Region	Connectedness rating (%)			Most connected herd	
			Average	Minimum	Maximum	Herd No	Region
PURELINE SWINE	1486	ON	8.8	1.8	16.4	11	ON
J. & R. PERREAULT	174	PQ	8.0	3.1	22.5	198	PQ
NORQUAY HOG FARM	18	MN	6.4	1.7	14.9	43	ON
FERME J.P. DION ET	198	PQ	6.2	2.4	22.5	174	PQ
VISTA VILLA FARMS LTD.	171	ON	6.2	1.3	16.4	1486	ON
THAMES BEND FARMS LTD.	43	ON	6	1.7	14.9	18	MN
CLARION HAMPSHIRE	11	ON	5.8	1.2	16.4	1486	ON
BLOOMSBURY FARMS LTD.	684	AB	5.4	1.1	47.2	1720	AB
PEAK SWINE GENETICS INC.	1720	AB	5.3	1.1	47.2	684	AB
BODMIN LIMITED	738	ON	5.1	1.3	13.5	1553	ON
QUALITY SWINE CO-OP	1553	ON	4.9	1.6	13.5	738	ON
NORQUAY DAUGHTER HERD	6019	MN	4.8	1.2	13.3	1486	ON
GENETIPORC	169	PQ	4.4	0.9	45.1	246	PQ
ELEVAGES C. & S.	269	PQ	4.3	2	11.6	174	PQ
JACK BRAAT	224	AB	4.2	0.7	14.4	1486	ON
GENETIPORC	246	PQ	4.2	0.7	45.1	169	PQ
LLOYD & SHIRLEY EVANS	1310	NS	3.9	1.2	10.3	171	ON
ELDALE SWINE LIMITED	1729	ON	3.7	0.9	10.3	1486	ON
STARDOBIE FARMS	232	ON	3.3	0.9	8.6	43	ON
FERME DENIS	189	PQ	3.1	1.3	8.5	174	PQ

Herd no. 1486 ( Pureline Swine ) had the highest average connectedness rating. This herd shares its genetics with 8 herds. Hence, it has high connectedness ratings with several herds. Most of the other herds have strong genetic links with one or few herds and almost no connections to most others herds. Hence, their averages are low.

Among Duroc herds (Appendix III), herd no. 252 (Ferme Frampton) has 22 boars that have been used in 15 other herds. Hence, it has the highest average of 6%. The highest connectedness rating (80%) was observed between herd no. 169 and 246, both belonging to Genetiporc. They had 19 common sires that had 99% progeny in one herd and 100% progeny in the other.

For the Yorkshire breed, sow productivity records were used. The ratings are given for herds contributing sow productivity data and having more than 10 records in the last contemporary group in 1996. Among Yorkshire herds (Appendix IV), the maximum connectedness (39%) was observed between herd no. 131 (Ferme Gaudreau) and herd no. 269 (Elevages C & S ) managed by the same owner.

Connectedness ratings were also calculated for contemporary groups within herds. They were higher than between herd ratings because of stronger genetic links within herds.

## **6 Strategies for increasing the level of connectedness**

### **6.1 Effect of using common sires on connectedness rating**

It is important to examine the behavior of the connectedness rating as genetic relationships between herds are changing, and to determine if it can be improved by the use of common sires.

Two moderately connected herds were selected as an example. One of them was in Ontario (Dietrich Farms Inc.) and the other in British Columbia (Dogwood Hybrid Pigs). The size of their latest contemporary groups for Duroc probe records were 67 and 75, respectively, representing the median group size for Duroc herds.

A hypothetical situation was constructed by replacing one sire in the Ontario herd with one sire in the B.C. herd. This simulates a situation whereby the herds would have used one common sire. Another example was created with two common sires, and so on. The results are given in Table 5.

Table 5: Effect of using common sires across two herds.

	Number of common sires			
	0	1	2	3
Connectedness rating	10	54	55	64
Standard error of difference				
Backfat (mm)	0.78	0.48	0.47	0.40
Age to 100 kg (days)	6.0	4.0	3.4	3.4

The results show that the connectedness increases with the use of common sires. The larger the number of common sires, the higher the connectedness rating. The standard error of the difference between herd effects also decreases, as expected. Including just one sire creates a very big change in the value of the connectedness rating and of the standard error. The gain is smaller for additional sires.



The common sires were used over several years, so that the magnitude of the changes in the connectedness rating might have been somewhat higher than in a real situation. Nevertheless, this example demonstrates how a change in the use of common sires can affect connectedness rating of two herds independently from any other effect.

## 6.2 Relationship between connectedness rating and proportion of progeny from common sires

In practice, there can be different numbers of common sires between herds and their use in each herd can differ. For example, one can mate a sire from another herd to a large or a small number of dams in one's own herd. Hence, the effect of the breeding structure of the herd on the degree of connectedness was examined.

A simple way of looking at this is to look at the relationship between the connectedness rating of two herds, the number of sires they have in common and the proportion of litters in each herd from these sires. This can be illustrated by a three dimensional diagram where the percentages of litters from common sires in each of the two herds are given on the two horizontal axis and the connectedness rating on the vertical axis (Figure 1). The figure shows that connectedness increases as the proportion of litters from the common sire increases. It is very high if 100% of the litters in both herds are from common sires. If the proportion of litters from the common sires is low in one of the herds, the other herd has to have a larger proportion of litters to reach a higher level of connectedness.

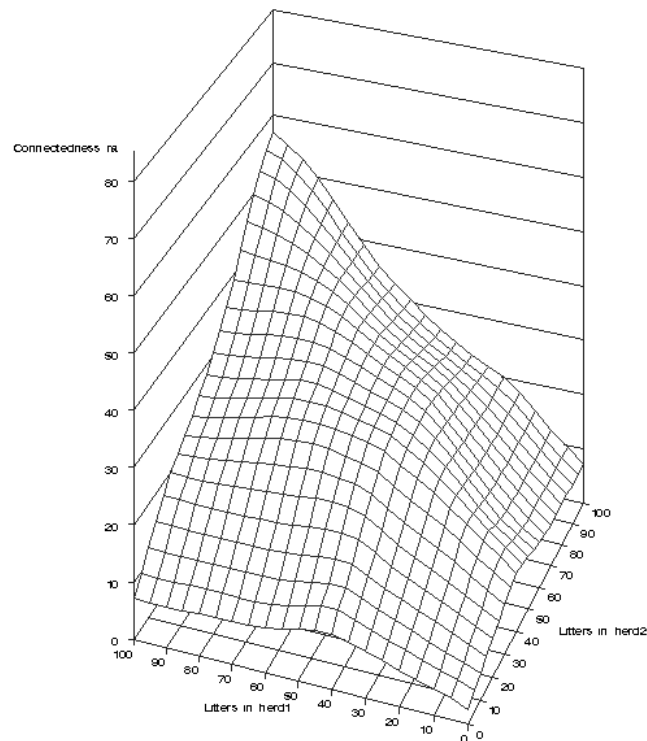


Fig. 1 Relationship between connectedness rating and percentage of litters from common sires in two Hampshire herds

The effect of choosing sires with a lower or higher percentage of litters in other herds is shown in Figure 2. The Figure shows that there are different optimum levels of connectedness ratings that can be achieved through the use of common sires depending upon the percentage of litters they have in other herds. If the sires have less than 25% of litters in another herd, the maximum level of connectedness rating possible is about 15% even if more than 50% of the litters in the original herd are produced from the common sires. However, if the sires have more than 75% of the litters in the other herd, the connectedness rating increases with the percentage of litters in the original herd.

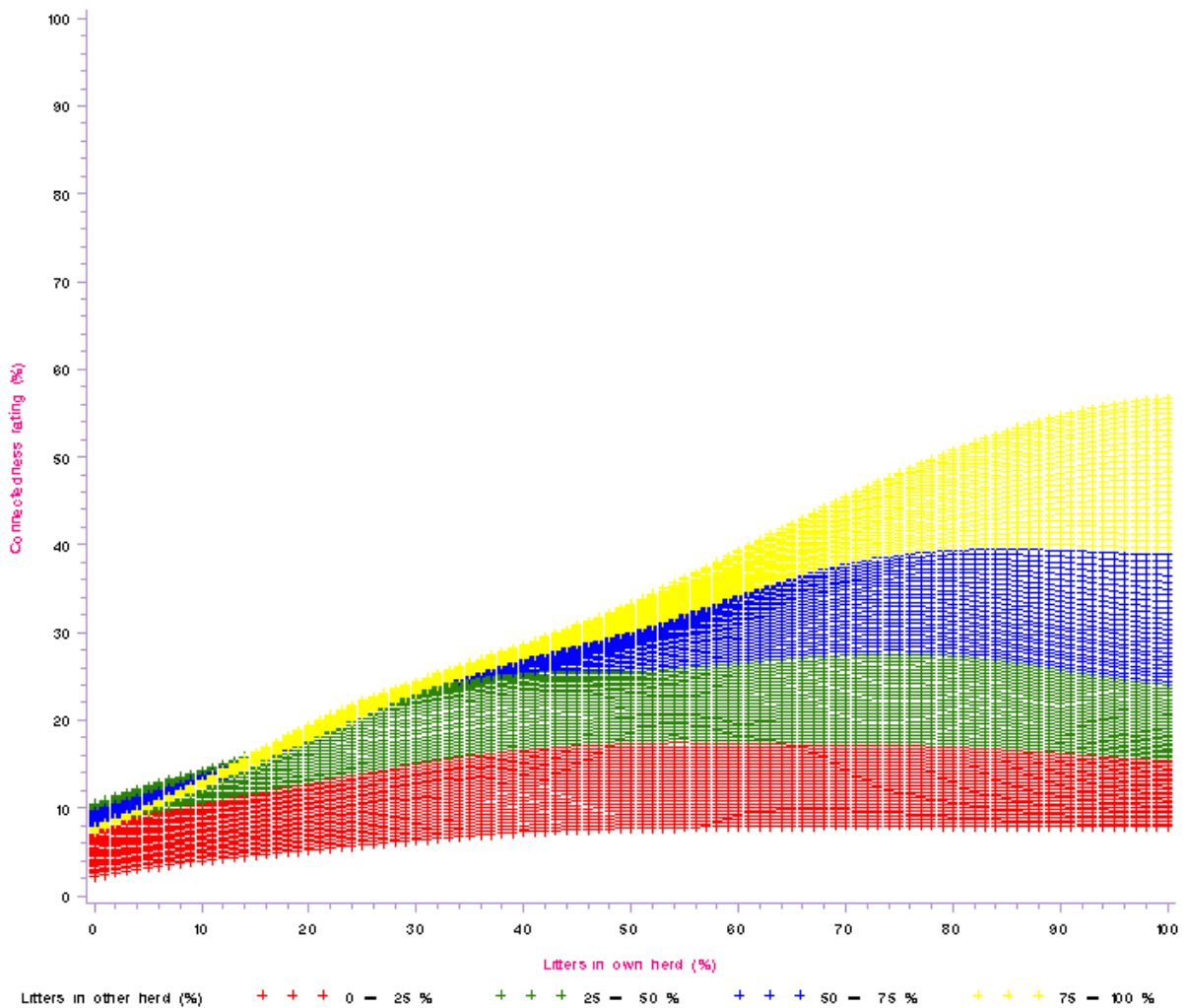


Fig. 2 Relationship between connectedness rating and percentage of litters in one's own herd for different percentages of litters in another herd

Even higher degrees of connectedness may be possible as a result of connections through common ancestors such as grand parents or as a result of placing pigs from the same litters in two contemporary groups. Figure 3 for Duroc herds shows that two herds have a connectedness rating as high as 80%. These herds belongs to the same company and pigs from the same litters are split across the two herds. Some herds have a high connectedness rating even if they have a low proportion of litters from common sires in the last contemporary group. Norquay Hog Farm and Norquay Daughter Farm have two common sires with 7% of the litters in one herd, and 21% of the litters in the other herd, but a connectedness rating of 29%. This is because the animals are related through other ancestors as well ( e.g. grand parents ).

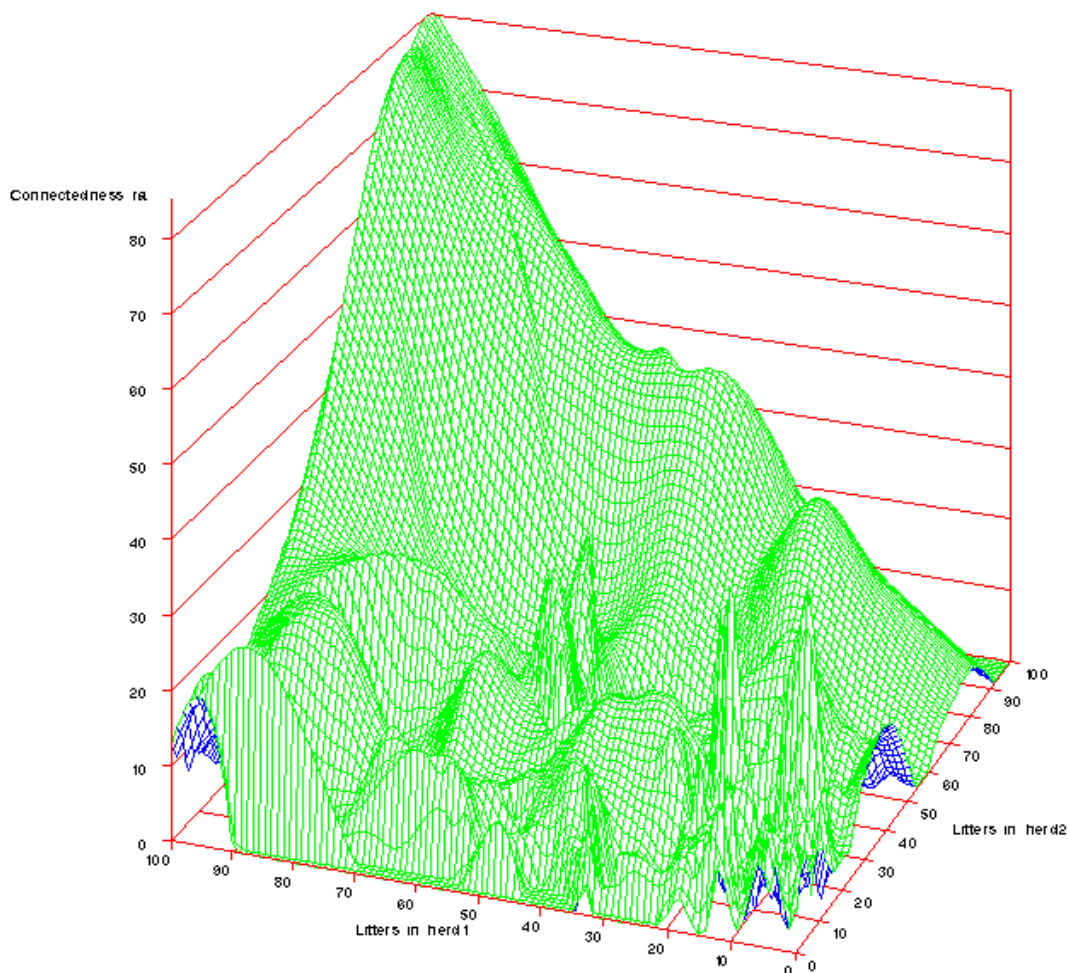


Fig. 3 Relationship between the percentage of litters from common sires and the connectedness rating of Duroc herds

7. **Relationship between the connectedness rating and the standard error of the difference between herd effects**

To develop guidelines for appropriate levels of connectedness among herds on the national program, it is important to know the effect of the connectedness rating on the standard error of comparison between EBVs in different herds.

As indicated earlier, the PEV of pairwise comparisons between all animals for all pairs of herds could be very difficult to compute. However, the PEV of the differences between estimates of herd effects is highly correlated with the average PEV of differences between EBVs. Hence, PEV of differences between herd effects were estimated for all pairs of herds. The square root of these estimates were taken as the standard error of prediction between herd estimates and their relationship with connectedness rating was studied.

There were 71 Duroc herds. Hence there were 2485 pairs of herds. The relationship between the connectedness ratings of these pairs and the SEP of difference between the herd effects is given in Figure 4.

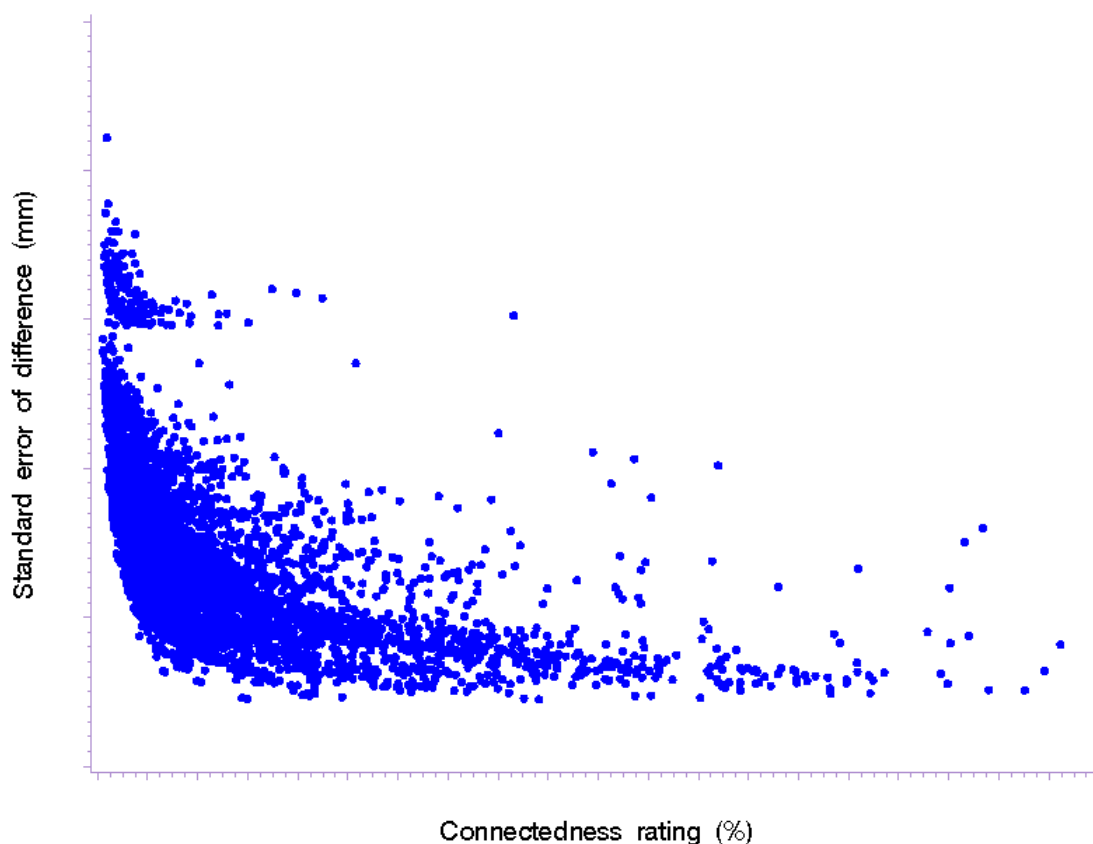


Fig. 4 Relationship between the connectedness rating and the standard error of differences between pairs of herd effects in backfat

The figure shows a nonlinear relationship between the two variables which could be expected since the SEP of differences between herd effects also depends on herd size and structure (equation (4)). Also as expected, the SEP decreases as the connectedness rating increases. The connectedness rating is very low for many pairs of herds. Their standard error of prediction ranges between 0.5 to 2.0 mm. If we convert these values to a confidence range (95%), the results show that one may expect an error ranging from  $\pm 1$  to  $\pm 4$  mm when comparing backfat EBVs of animals across herds. The error is specially large when connectedness is below 5%. There is little gain in accuracy when the connectedness rating increases above that point.

The relationship between the average connectedness rating of each herd with the average standard error of differences with all other herds is shown in Figure 5. The trend is similar to the pairwise comparisons except that the extreme values are averaged out. The average values are lower than for pairwise comparisons. There are many herds with low average connectedness ratings. The average standard error decreases when average connectedness rating increases, therefore the higher the rating the more accurate the comparisons. It seems important to have at least a 3% average connectedness rating for reasonably accurate comparisons of backfat and age EBVs.

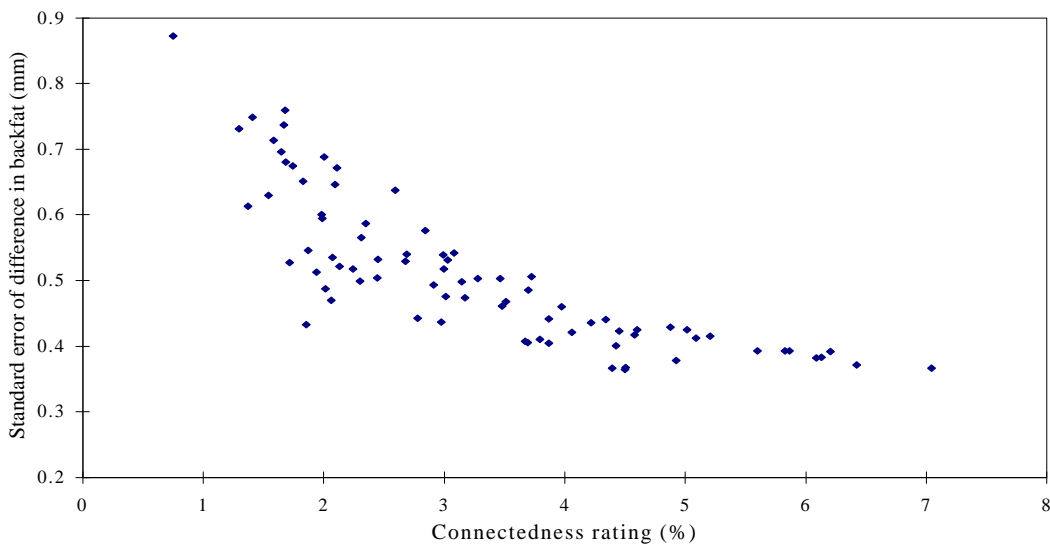


Fig. 5 Relationship between the average connectedness rating and the average standard error of differences between herd effects in backfat

The SEP of differences is affected by the size of the herds and their structure. As shown in Figure 6, the SEP of herd effects depends strongly on contemporary group size. The SEP is very large when contemporary group size is low, especially when it is below 10.

For backfat and age, an average connectedness rating of more than 3% and a minimum management group size of 10 animals would be required to ensure that the EBVs of a herd on the program can be accurately compared to that of other herds. Essentially, the SEP tends to plateau when the herd size or connectedness rating become sufficiently large.

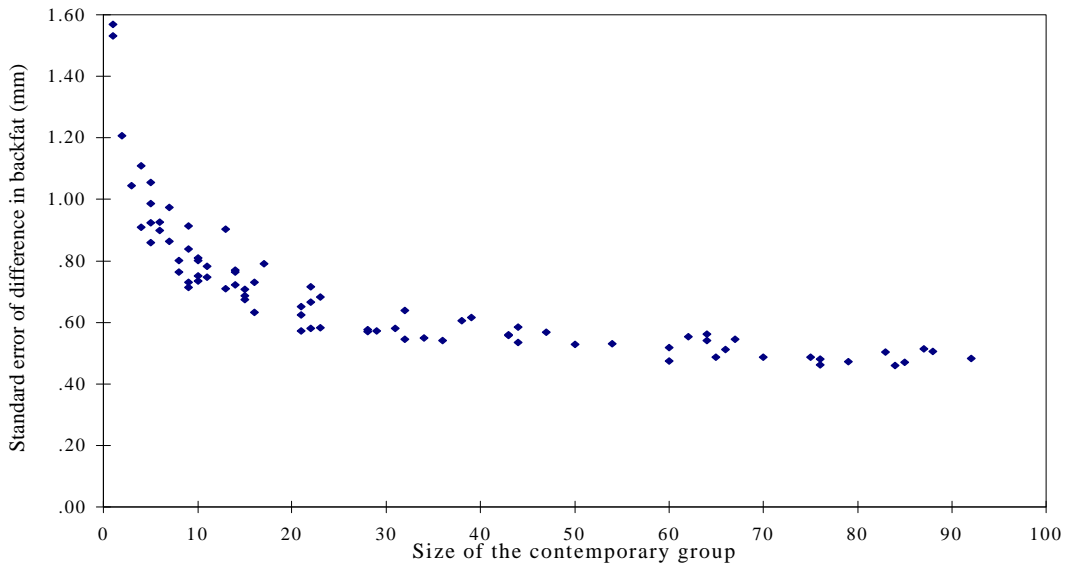


Fig. 6 Relationship between the contemporary group size and the average standard error of differences between group effects in backfat.

For litter size, connectedness ratings are lower and minimum connectedness ratings of 1.5% is recommended ( Figures 7 and 8 ).

## 8. Conclusions

The results show that the correlation between estimates of herd effects is a useful estimate of connectedness as seen from the actual data. This correlation accounts for the relationship between animals from each herd and their relationships to all other herds included in the evaluation.

Most herds are well connected to one or few other herds. However, the degree connectedness to most other herds in the national program is low. If this trend continues, the accuracy of the comparisons across herds will be adversely affected. Therefore, the average connectedness of all herds should be brought to a minimum level.

A simple approach to increase connectedness is to use common sires with other herds. A connectedness rating of 3 to 18% can be achieved with another herd by using one common sire. If the sire selected is a widely used A.I. sire, connections can be improved simultaneously with several herds.

## **9. References**

Kennedy, B. W. and D. Trus (1993) Considerations on genetic connectedness between management units under an animal model, *J. Anim. Sci.* 71: 2341-2352

Banos, G. and R. A. Cady (1988) Genetic relationship between the United States and Canadian Holstein bull populations. *J. Dairy Sci* 71:1346-1354

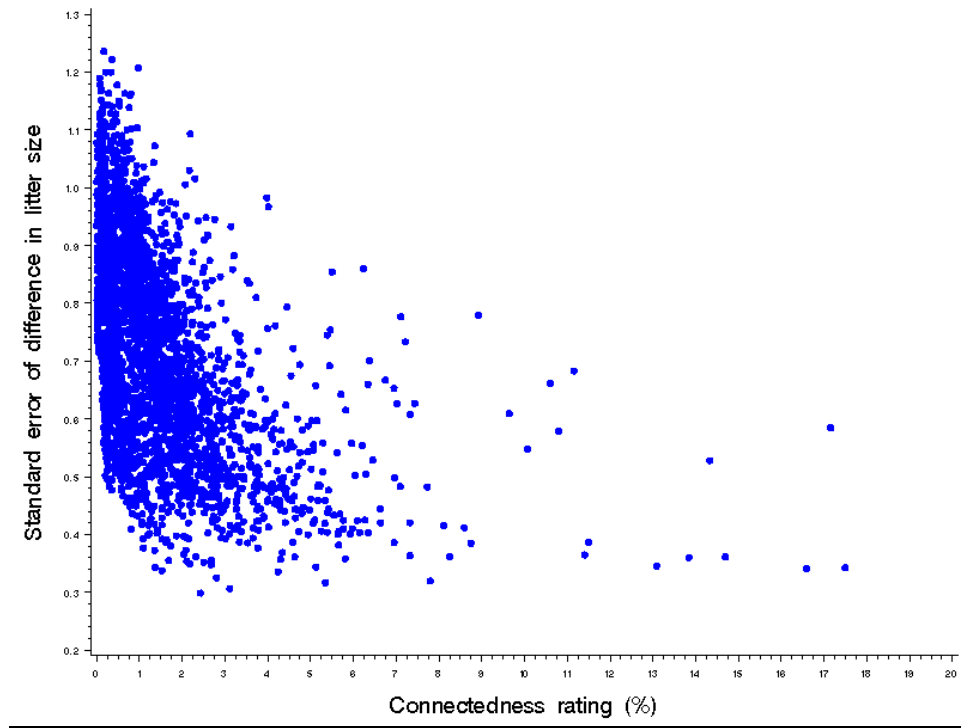


Fig. 7 Relationship between the connectedness rating and the standard error of differences between pairs of herd effects in litter size

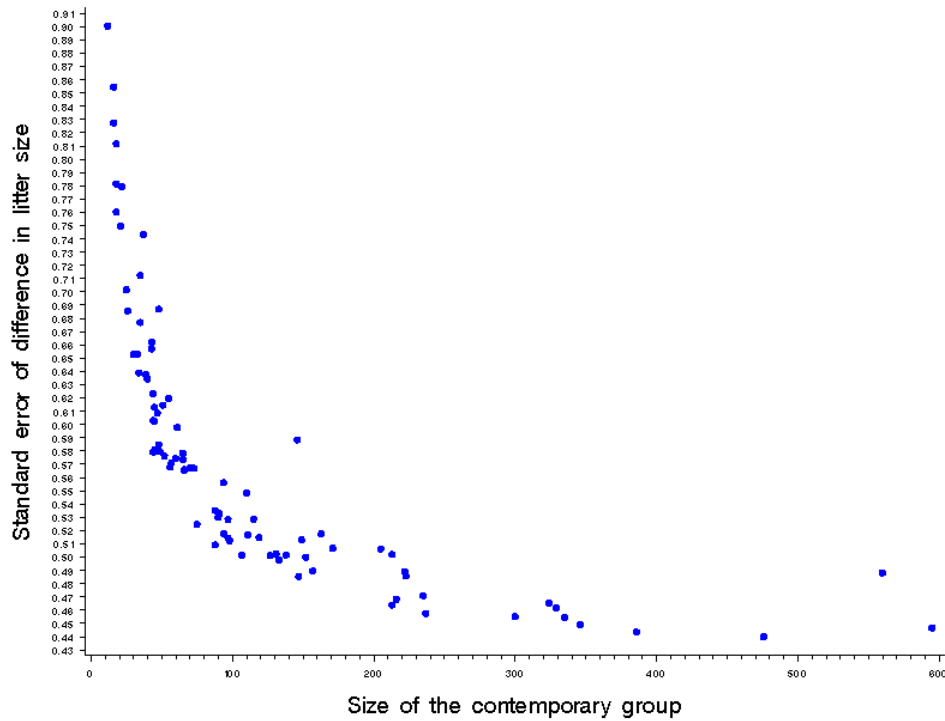


Fig. 8 Relationship between the contemporary group size and the average standard error of differences between herd effects in litter size



## **10. Recommendations**

### **1) Minimum requirements**

- contemporary group size should be at least 10 pigs or sows from at least 3 different sires;
- the average connectedness rating (CR) of each herd should be 3% or more for probe records and 1.5% or more for litter records;
- if substantial importation of genetics from a herd is being considered, the minimum CR with that herd should be 5%.

### **2) Methods for increasing connectedness ratings**

- if the average CR is lower than the recommended values of 3% or 1.5 %, use sires from the list of AI boars with high CR (Appendix V), or use sires from herds with high CR ratings;
- using one well-connected sire from the high CR sire list with a sufficient proportion of progeny in your herd (15%) will usually be enough to reach the 3% minimum level;
- if you use sires from herds with high CR ratings, these sires should have a sufficient proportion of progeny (15%) in both the herd of origin and your herd.

### **3) Creation of a genetic pool**

- in order to have well-connected boars of sufficient genetic merit, it is necessary that breeders participating in the program make some of their superior genetics available;
- it is therefore recommended that a national genetic pool of AI boars be established with cooperating AI Centres. Participation in the pool would be voluntary but only those who participate would be allowed to draw from the pool. The participating breeders would make available for potential use in the pool genetics from the top 20% of their herd each year (e.g. the AI Centres would have a choice of boars from a certain number of litters designated by the breeder). Some restrictions on the use of semen from specific boars (i.e. an individual herd may only use so many inseminations from specific boars) as well as special royalty considerations to reward breeders that provide top genetics could also be considered;
- these recommendations allow each breeder to keep their very best genetics within their herd (the top 5 to 10%), while giving participants to the pool access to a broader range of top genetics. Breeders participating in the pool become part of a

super-nucleus structure which allows them to compete effectively at the international level;

- to establish this national pool of AI boars, the following steps could be taken:
  - a) regional centres and AI Centres could prepare a proposal with technical guidance from CCSI;
  - b) the proposal could be presented to breeders on the program;
  - c) if there is enough support, those breeders that are interested could enter into an agreement with AI Centres which would describe the conditions of participation and access to the pool.

**Appendix I: Examples to illustrate the method of estimating connectedness and its effect on comparison of EBVs**

**Case - 1** : Unrelated individuals each with a single backfat record.

	Herd $h_1$	Herd $h_2$	
animals	$a_1$ $a_2$	$a_3$ $a_4$	unrelated

A simple animal model would be:

$$Y_{ij} = h_i + a_{ij} + e_{ij}$$

where  $Y_{ij}$  is the observation on the  $ij^{\text{th}}$  animal ( $a_{ij}$ ) made in the  $i^{\text{th}}$  herd ( $h_i$ ), and  $e_{ij}$  is the random residual effect. Using this information, the coefficient matrix of the mixed model equations was written considering the heritability ( $h^2$ ) of backfat as 0.52 and the residual variance ( $\sigma_e^2$ ) as 1.69  $\text{mm}^2$  as in national genetic evaluations. A direct inverse was calculated and the following inverse elements (the prediction error variance matrix) were obtained:

	$h_1$	$h_2$	$a_1$	$a_2$	$a_3$	$a_4$
$h_1$	1.76	0	-0.92	-0.92	0	0
$h_2$	0	1.76	0	0	-0.92	-0.92
$a_1$	-0.92	0	1.355	0.476	0	0
$a_2$	-0.92	0	0.476	1.355	0	0
$a_3$	0	-0.92	0	0	1.355	0.476
$a_4$	0	-0.92	0	0	0.476	1.355

Here, the  $\text{cov}(\hat{h}_{12}) = 0$  and  $\text{var}(\hat{h}_1) = \text{var}(\hat{h}_2) = 1.76$ .

$$\text{CR}_{12} = \frac{\text{cov}(\hat{h}_{12})}{\sqrt{\text{Var}(\hat{h}_1)\text{Var}(\hat{h}_2)}} \times 100 = \frac{0}{\sqrt{(1.76)(1.76)}} \times 100 = 0\% , \text{ and}$$

$$\text{var}(\hat{h}_i - \hat{h}_j) = \text{var}(\hat{h}_1) + \text{var}(\hat{h}_2) - 2 \text{cov}(\hat{h}_{12}) = 1.76 + 1.76 - 2(0) = 3.52 \text{ mm}^2$$

The PEV of the EBV of each animal is 1.355. Hence, the repeatability is:  $(1 - \text{PEV}/\sigma_a^2) \times 100 = 43\%$  assuming a genetic variance of 2.36  $\text{mm}^2$ .

The PEV of the difference between the EBVs of animals within herd is:

$$\text{PEV}(\hat{a}_1 - \hat{a}_2) = \text{var}(\hat{a}_1) + \text{var}(\hat{a}_2) - 2 \text{cov}(\hat{a}_{12}) = 1.355 + 1.355 - 2(0.476) = 1.758 \text{ mm}^2$$

The PEV of the difference between animals from different herds is:

$$PEV(\hat{a}_1 - \hat{a}_3) = 1.355 + 1.355 - 2(0) = 2.71 \text{ mm}^2$$

This is higher than the within herd PEV of EBVs because there is no covariance between unrelated animals from different herds.

**Case - 2** - Animals related within herd but unrelated across herds.

	Herd 1		Herd 2		
animal	a <sub>1</sub>		a <sub>3</sub>		Full sibs within herds
	a <sub>2</sub>		a <sub>4</sub>		

Here, a<sub>1</sub> and a<sub>2</sub> are a pair of full-sibs and a<sub>3</sub> and a<sub>4</sub> are another pair of full-sibs. The PEV matrix part for the herds and animals is given below.

	h <sub>1</sub>	h <sub>2</sub>	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>	a <sub>4</sub>
h <sub>1</sub>	2.148	0	-1.3	-1.3	0	0
h <sub>2</sub>	0	2.148	0	0	-1.3	-1.3
a <sub>1</sub>	-1.3	0	1.6	1.006	0	0
a <sub>2</sub>	-1.3	0	1.006	1.6	0	0
a <sub>3</sub>	0	-1.3	0	0	1.6	1.006
a <sub>4</sub>	0	-1.3	0	0	1.006	1.6

$$CR_{12} = \frac{\text{cov}(\hat{h}_{12})}{\sqrt{\text{Var}(\hat{h}_1)\text{Var}(\hat{h}_2)}} \times 100 = \frac{0}{\sqrt{(2.148)(2.148)}} \times 100 = 0\% , \text{ and}$$

$$PEV(\hat{h}_1 - \hat{h}_2) = \text{var}(\hat{h}_1) + \text{var}(\hat{h}_2) - 2 \text{cov}(\hat{h}_{12}) = 2.148 + 2.148 - 2(0) = 4.296 \text{ mm}^2$$

The PEV of the EBV of each animal is 1.6. Hence, the repeatability is:  $(1 - 1.6/2.36) \times 100 = 32\%$

The PEV of the difference between the EBVs of animals within herd is:

$$PEV(\hat{a}_1 - \hat{a}_2) = \text{var}() + \text{var}() - 2 \text{cov}() = 1.6 + 1.6 - 2(1.006) = 1.188 \text{ mm}^2$$

The PEV of the difference between animals from different herds is:

$$PEV(\hat{a}_1 - \hat{a}_3) = 1.6 + 1.6 - 2(0) = 3.2 \text{ mm}^2$$

The comparison of cases 1 and 2 reveals that if the animals are related within herd, the PEV of the difference between animals within herd is lower and that of animals across herds is higher.

**Case - 3** Animals unrelated within herd but related across herds.

	herd h <sub>1</sub>	herd h <sub>2</sub>	
Animals	a <sub>1</sub>	a <sub>3</sub>	Full sibs
	a <sub>2</sub>	a <sub>4</sub>	across herds

Here, a<sub>1</sub> and a<sub>3</sub> are a pair of full-sibs across herds and a<sub>2</sub> and a<sub>4</sub> are another pair of full-sibs. The PEV matrix part for the herds and animals is given below.

	h <sub>1</sub>	h <sub>2</sub>	a <sub>1</sub>	a <sub>2</sub>	a <sub>3</sub>	a <sub>4</sub>
h <sub>1</sub>	1.725	0.423	-0.88	-0.88	-0.42	-0.42
h <sub>2</sub>	0.423	1.725	-0.42	-0.42	-0.88	-0.88
a <sub>1</sub>	-0.88	-0.42	1.285	0.475	0.53	0.315
a <sub>2</sub>	-0.88	-0.42	0.475	1.285	0.315	0.53
a <sub>3</sub>	-0.42	-0.88	0.53	0.315	1.285	0.475
a <sub>4</sub>	-0.42	-0.88	0.315	0.53	0.475	1.285

$$CR_{12} = \frac{\text{cov}(\hat{h}_{12})}{\sqrt{\text{Var}(\hat{h}_1)\text{Var}(\hat{h}_2)}} \times 100 = \frac{0.423}{\sqrt{(1.725)(1.725)}} \times 100 = 25\% , \text{ and}$$

$$\text{PEV}(\hat{h}_1 - \hat{h}_2) = \text{var}(\hat{h}_1) + \text{var}(\hat{h}_2) - 2 \text{cov}(\hat{h}_{12}) = 1.725 + 1.725 - 2(0.423) = 2.604 \text{ mm}^2$$

The PEV of the EBV of each animal is 1.285. Hence, the repeatability is:  
 $(1 - 1.285/2.36) \times 100 = 46\%$

The PEV of the difference between EBVs of unrelated animals within herd is:

$$\text{PEV}(\hat{a}_1 - \hat{a}_2) = 1.285 + 1.285 - 2(0.475) = 1.62 \text{ mm}^2$$

The PEV of the difference between related animals from different herds is:

$$\text{PEV}(\hat{a}_1 - \hat{a}_3) = 1.285 + 1.285 - 2(0.53) = .151 \text{ mm}^2$$

The PEV of the difference between unrelated animals from different herds is:

$$\text{PEV}(\hat{a}_1 - \hat{a}_4) = 1.285 + 1.285 - 2(0.315) = 1.94 \text{ mm}^2$$

Thus, the PEV of the difference between herds decreases as the connectedness between herds increases.