

8-2010

Comparison Of Connectedness Measures And Changes In Connectedness Of The U.S. Duroc Population

N. Soga

Iowa State University

Matthew L. Spangler

University of Nebraska - Lincoln, mspangler2@unl.edu

C. R. Schwab

National Swine Registry

P. J. Berger

Iowa State University

T. J. Baas

Iowa State University

Follow this and additional works at: <http://digitalcommons.unl.edu/animalscifacpub>

Soga, N.; Spangler, Matthew L.; Schwab, C. R.; Berger, P. J.; and Baas, T. J., "Comparison Of Connectedness Measures And Changes In Connectedness Of The U.S. Duroc Population" (2010). *Faculty Papers and Publications in Animal Science*. 801.
<http://digitalcommons.unl.edu/animalscifacpub/801>

This Article is brought to you for free and open access by the Animal Science Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Papers and Publications in Animal Science by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Comparison Of Connectedness Measures And Changes In Connectedness Of The U.S. Duroc Population

N. Soga^{*}, M. L. Spangler[†], C. R. Schwab[‡], P. J. Berger^{*}, and T. J. Baas^{*}

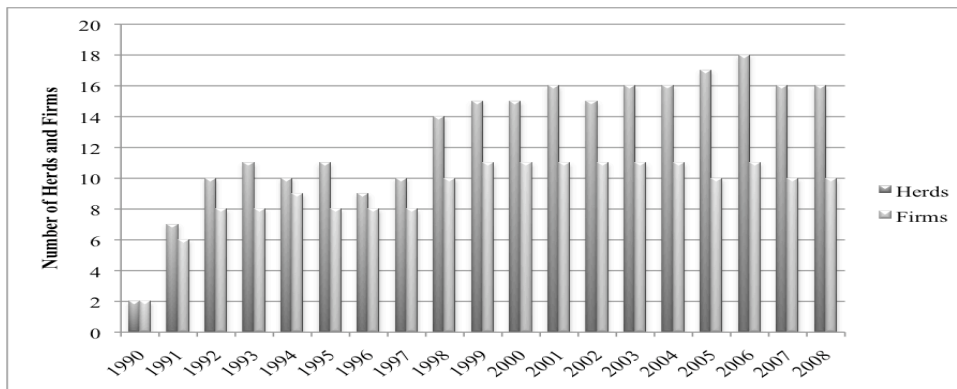
Introduction

Low levels of connectedness can create problems in uncoupling genetic and environmental influences on phenotypes. This can result in bias in prediction error variance and covariance, decreasing the accuracy of comparison of EBV across herds (Smith and Banos, 1991; Kennedy and Trus, 1993). The main objective of this study was to compare connectedness statistics published in the literature and determine which measure is most practical or suitable for estimating the risk associated with EBV comparisons in an across-herd analysis and to determine changes in connectedness in the U.S. Duroc population over time.

Material and methods

Data from the National Swine Registry (NSR) Swine Testing and Genetic Evaluation System (STAGES) from Duroc litters born from 1990 to 2008 were used including 156,101 records from 160,325 animals representing 33,210 litters, 21 herds, and 11 firms. The numbers of herds within breeder-nucleus firms utilized in the analysis are presented by year (Figure 1) where a firm is defined as an entity that controls management decisions over multiple. The herds represented in the current study are those that were still in existence in 2008. Four common connectedness statistics were estimated: prediction error variance of all pair-wise EBV differences between the animals in those herds (PEVD) (Kennedy and Trus, 1993), connectedness correlation (R) (Mathur et al., 2002), connectedness rating (CR) (Lewis et al., 1999), and common sire percentage (CS%).

Figure 1. Number of herds and firms in the U.S. purebred Duroc population recorded each year by the National Swine Registry's Swine Testing and Genetic Evaluation System (STAGES) from 1990 to 2008



Data for ultrasonically measured backfat were analyzed according to the following single-trait model:

$$Y_{ijkm} = u + cg_i + sex_j + litter_k + a_{ijkm} + wt_m + wt_m^2 + e_{ijkm}$$

where Y_{ijkm} = the trait measured on pig m in contemporary group i of sex j in litter k ; u = mean; cg_i = fixed effect of contemporary group i ; sex_j = fixed effect of sex j ; $litter_k$ = random effect of litter k ; a_{ijkm} = effect of animal m assumed random; wt_m and wt_m^2 = linear and quadratic effects of the offset weight of pig m ; e_{ijkm} = random residual error. Variance and covariance components were estimated using MTDFREML (Boldman

^{*} Iowa State University, Department of Animal Science, Ames, IA, 50011, USA.

[†] University of Nebraska-Lincoln, Department of Animal Science, Lincoln, NE, 68583, USA.

[‡] National Swine Registry, West Lafayette, IN, 47966, USA.

et al.,1995). Progeny were grouped by birth year and herd to define groups under similar common environments (Herd-Year).

Prediction error variance of differences between the animals (PEVD). PEVD was calculated as:

$$PEVD = PEV(\mathbf{u}_i - \mathbf{u}_j) = \sum PEV(\mathbf{u}_i) / R_i + \sum PEV(\mathbf{u}_j) / R_j - 2 PEC(\sum \mathbf{u}_i, \sum \mathbf{u}_j) / (R_i + R_j).$$

Individual EBV and their prediction error variances were calculated and averaged by Herd-Year:

$$\sum PEV(\mathbf{u}_i) / R_i,$$

where R_i is the number of total progeny in Herd-Year i .

Prediction error covariance for genetic group i and j was calculated by calculating contrasts of the sum of the individual EBV for Herd-Year i and j and averaged by the number of progeny:

$$PEC(\sum \mathbf{u}_i, \sum \mathbf{u}_j) / (R_i + R_j),$$

where R_i and R_j are the number of total progeny in herd i and herd j , respectively.

Connectedness correlation was calculated by the following: prediction error variance ($PEV(\sum \mathbf{u}_i)$) and covariance ($PEC(\sum \mathbf{u}_i, \sum \mathbf{u}_j)$) for Herd-Year i and j were estimated by calculating contrasts of the sum of individual BVs for each genetic group such that

$$R = PEC(\sum \mathbf{u}_i, \sum \mathbf{u}_j) / \sqrt{PEV(\sum \mathbf{u}_i) * PEV(\sum \mathbf{u}_j)}.$$

Connectedness rating (CR) was calculated by the following: variance of estimation of each Herd-Year effect ($Var(\sum \mathbf{g}_i)$) and covariance ($Cov(\sum \mathbf{g}_i, \sum \mathbf{g}_j)$) for Herd-Year i and j were estimated by calculating contrasts of the sum of contemporary group effects in each Herd-Year. Then,

$$CR = Cov(\sum \mathbf{g}_i, \sum \mathbf{g}_j) / \sqrt{Var(\sum \mathbf{g}_i) * Var(\sum \mathbf{g}_j)}.$$

Common sire percentage (CS%) was calculated by the following:

$$CS\% = (Pc_{(i)} * Pc_{(j)}) / (R_i + R_j)$$

for each pair of Herd-Year.

Results and discussion

Correlations among 3 connectedness statistics (CR, R, CS%), PEVD, and the absolute value of the genetic mean difference between herds (GMD) are presented in Table 1. High correlation estimates were obtained among the different measures of connectedness in the current data. Values for the relationship between PEVD and the 3 connectedness statistics were moderately negative (-0.34 to -0.24). Computationally, CS% is significantly less demanding when compared to R, whereas R is similar to that of PEVD, and CR is intermediate to R and CS%. These correlations were lower than previous studies (Kuehn et al., 2009), likely due to the unbalanced nature of the dataset used in the present study. Correlations between GMD and CR, R, and CS% were similar (-0.30 to -0.29). The correlation between GMD and PEVD was 0.11. This estimate was also smaller than that reported by Kennedy and Trus (1993).

Results from the current study suggest that pairs of herds within the same firm tend to have higher connectedness and lower PEVD than pairs of herds from different firms. This may be explained by the fact that, due to bio-security concerns and cost, sires are more commonly used across herds within a firm when compared to across firms, and herds within a firm generally have similar selection schemes, which should decrease the GMD. Of the 3 connectedness statistics, CS% is more variable than R and CR. Most of the outliers were pairs of genetic groups that contained less than 100 pigs. From these data, if average herd size is greater than 100, CS% can be used as an accurate indicator of connectedness.

For the 3 connectedness statistics (CR, R, and CS %), when levels of connectedness become greater than 0.1, PEVD was suppressed to the level of base line prediction error variance deviation (BL_{pevd}). Therefore, 0.1 can be considered as a criterion for non-biased levels of connectedness. These criteria are approximately equal to those reported in previous studies (Kuehn et al., 2008; Mathur et al., 2002). When data are unbiased ($PEVD = BL_{pevd}$), the accuracy associated with comparison of EBV is higher than when data are biased (high PEVD; low connectedness). For backfat, the accuracy associated with comparisons of EBV across herds decreased from 77.6% (unbiased) to 54.6% (biased).

Table 2. Correlations between GMD, PEVD, CS%, R, and CR from 1990 to 2008^a

	GMD	PEVD	CS%	R	CR
GMD	-----				
PEVD	0.115	-----			
CS%	-0.301	-0.244	-----		
R	-0.294	-0.341	0.968	-----	
CR	-0.296	-0.313	0.964	0.989	-----

^a GMD= Genetic mean difference; PEVD= Prediction error variance difference; CS%= Common sire percentage; R= Connectedness correlation; CR= Connectedness rating.

The PEVD decreased until 1997 and then slightly increased or was stable for the largest 3 firms. The same tendency can be seen for the mean of all herds (Figure 2). The R, CR, and CS% for all herds increased until 2003, and then decreased. The largest changes were in R (Figure 3) and CS%, reaching a maximum value of 0.07 and 0.05 in 2003, respectively, and then decreasing to 0.04 and 0.01, respectively. The GMD decreased until 1997 and then slightly increased or was stable for the largest 3 firms and the same tendency can be seen for the mean of all herds. GMD reached a maximum value in 1990 (0.18) and declined to 0.02 in 2001 and then increased to 0.05 in 2008. The current levels of connectedness within the U.S. purebred Duroc breed, when herd sizes are small, are lower than each criteria associated with increased risk of bias (CR: 3-5%, R: 5-10%, CS%: 5-10%) (Mathur et al., 2002; Kuehn et al., 2008). The changes in connectedness overtime reflect changes in the structure of the US Duroc population as the number of firms decreases and the size of the remaining firms increases in total number of pigs.

Figure 2. Trend in the mean and standard deviation of change of prediction error variance of all pair-wise estimated BV differences (PEVD) for all herds in the U.S. purebred Duroc population from the National Swine Registry Swine Testing and Genetic Evaluation System (STAGES) from 1990 to 2008.

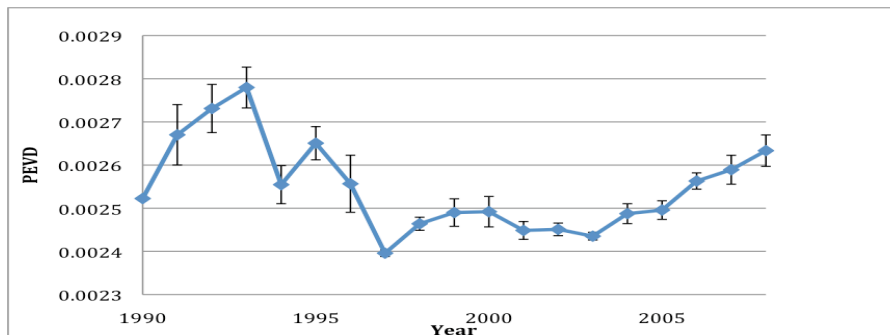
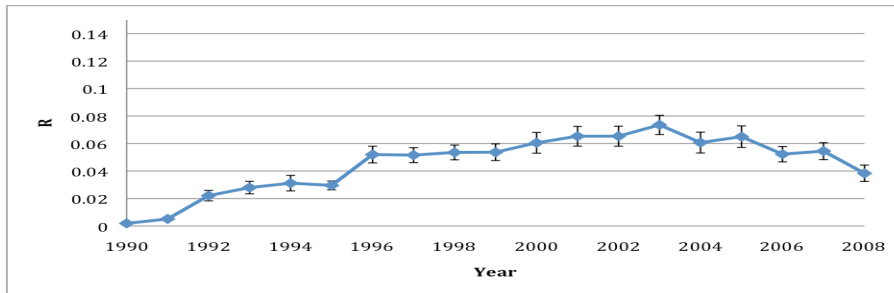


Figure 3. Trend in the mean and standard deviation of change of connectedness correlation (R) for all herds in the U.S. purebred Duroc population from the National Swine Registry Swine Testing and Genetic Evaluation System (STAGES) from 1990 to 2008.



Conclusion

Results of this study indicate that increasing values for connectedness decreases bias (low PEVD) of comparisons of EBV across herds. These 3 connectedness statistics were highly and positively correlated. Correlations between these 3 connectedness statistics and bias were largest for R, and smallest for CS%. Even though R was the most accurate indicator of connectedness, this statistic was computationally demanding to obtain. Therefore, CR was most suitable as an indication of connectedness, due to its ease of computation. The CS% can also be an indication of connectedness, but only when pairs of herds are relatively large (>100 animals). When these connectedness statistics become lower than approximately 10%, the risk of bias significantly increases. In order to investigate the effect of heritability and record density in multiple trait models, analysis using additional traits of interest such as maternal traits that have lower heritability and are sparsely recorded will be required. Further studies in this U.S. Duroc population should evaluate the number of common sires needed for maternal traits where records are obtained only in 1 sex and later in life, as well as simulation studies that examine the effectiveness of recently implemented sire sampling programs. Results suggest there is a need to establish increased levels of connectedness within U.S. purebred nucleus herds for more accurate comparison of EBV through effective exchange of common reference sires across nucleus herds.

References

- Boldman, K.G., Kriese, L.A., Van Vleck, L.D. *et al.* (1995). U.S. Dept. of Agriculture, Agricultural Research Service. Univ. of Nebraska, Lincoln, NE.
- Kennedy, B.W., and Trus, D. (1993). *J. Anim. Sci.* 71:2341-2352.
- Kuehn, L.A., Lewis, R.M., and Notter, D.R. (2008). *Genet. Sel. Evol.* 39:225-247.
- Kuehn, L.A., Lewis, R.M., and Notter, D.R. (2009). *J. Anim. Sci.* 87:507-515.
- Lewis, R.M., Crump, R.E., Simm, G. *et al.* (1999). In *Proc. Brit. Soc. Anim. Sci.*, page 121.
- Mathur, P.K., Sullivan, B.P., and Chesnais, J.P. (2002). In *Proc. 7th WCGALP*, volume 32, pages 545-548.
- Smith, C., and Banos, G. (1991). *J. Anim. Sci.* 69:2387-2394.