

Evaluation of connectedness between herds for three pig breeds in China

C. Y. Sun^{1,2}, C. K. Wang^{1,2}, Y. C. Wang^{1,2}, Y. Zhang^{1,2} and Q. Zhang^{1,2†}

¹The Key Laboratory of Animal Genetics and Breeding of the Ministry of Agriculture, China Agricultural University, Beijing 100094, China; ²The State Key Laboratory for Agrobiotechnology, China Agricultural University, Beijing 100094, China

(Received 6 April 2008; Accepted 6 November 2008; First published online 23 December 2008)

Using data on age at 100 kg of three pig breeds (Large White, Landrace and Duroc) the connectedness between herds in China was evaluated by the connectedness rating (CR) method. The results show that most herds in China have low average CR (0–3.59%) with other herds. In Large White, of the 36 herds analyzed there are 20 herds, which are connected with at least one other herd. In Landrace and Duroc, of the 27 and 18 herds analyzed, only four and five herds, respectively, were found having connectedness with other herds. Generally, the connectedness exists only among two or few herds in the same region. A certain degree of connectedness between herds is a precondition for cross-herd genetic evaluation. A national or regional cross-herd genetic evaluation is not practicable at the present time in China. More intense efforts are needed to establish and enhance the connectedness between herds by means of extensively using artificial insemination (AI) in the swine industry in China.

Keywords: connectedness, cross-herd genetic evaluation, joint breeding system

Introduction

The BLUP genetic evaluation method enables the evaluation of animals of different herds jointly and allows the comparison of estimated breeding values (EBVs) of animals in different herds. However, the accuracy of such comparison depends on the degrees of connectedness between herds. The higher the connectedness, the more accurate the comparisons. Therefore, it is important to measure the degree of connectedness among herds, and bring it to a level that allows across-herd comparisons of EBVs with reasonable accuracy.

Several methods have been developed to measure the degree of connectedness. Foulley *et al.* (1990 and 1992) proposed a connectedness index (CI). Laloë (1993) introduced the concept of generalized coefficient of determination (CD). Fries (1998) suggested measuring connectedness using the number of direct genetic links between herds due to common sires and dams, which, although very simple, generally underestimates the connectedness because it ignores the indirect links and the links due to common environment. Since the objective of measuring connectedness is to obtain an indication of the accuracy as well as the bias of comparisons between EBVs in different herds, Kennedy and Trus (1993) suggested that the most appropriate measure of connectedness would be the average

prediction error variance (PEV) of differences in EBVs between animals in different management units. Laloë *et al.* (1996) compared PEV with CI and CD and concluded that CD combines aspects of genetic variability and PEV, and is thus a method of choice to assess the connectedness. Kuehn *et al.* (2007) also argued that CD was a perfect indicator of potential bias remaining when comparing individuals in separate units. However, if the primary objective of measuring connectedness 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 that assesses only the accuracy of such comparisons would be most appropriate. Therefore, PEV was often chosen as the standard against which to evaluate all other methods (e.g. Mathur *et al.*, 2002; Roso *et al.*, 2002; Huisman *et al.*, 2006). The calculation of PEV (and CD) requires the inverse corresponding to the animal effects in the coefficient matrix of the mixed model equations (MME). This matrix is usually huge and its inverse is very difficult to be calculated even with the most modern computer. A simulation study by Kennedy and Trus (1993) suggested that the PEV of the differences between herds (PEVD) had a high correlation (0.995) with average PEV of differences in EBVs between animals, and can be used as an effective measure of the degree of connectedness between herds. However, the computational problem remains for large populations with many management units. In addition,

† E-mail: qzhang@cau.edu.cn

PEVD depends on the size and structure of herds. Mathur *et al.* (1998 and 2002) proposed using connectedness rating (CR), expressed as the correlation between the estimates of herd effects, to approximately estimate PEV. The authors proved CR is strongly related to PEV and less depends on herd size and variation than PEVD, and thus can be used to estimate connectedness. This method has been used routinely in the national swine genetic evaluation in Canada.

In China, a national swine improvement program was initiated in 1998 (Fu, 2005) with the main aim of setting up a national joint swine breeding system. The key point of the system is cross-herd genetic evaluation. The connectedness among herds has to be taken into account in this program. The objective of this study is to estimate the degree of connectedness among pig herds in China using the CR method of Mathur *et al.* (1998).

Material and methods

Data

Three data sets, which consist of records on age at 100 kg of three pig breeds from different herds, were used to study the connectedness between herds. The herds included in the data sets represent the major pig breeding farms in China. The pigs producing these records were born between 1998 and 2005. The numbers of animals and herds in each data set are shown in Table 1.

Model

The data sets were analyzed for each breed using the following linear animal model:

$$y = Xh + Za + Tp + e, \tag{1}$$

where *y* is the vector of records on age at 100 kg, *h* is the vector of herd effects, *a* is the vector of random animal additive genetic effects, *p* is the vector of random litter effects, and *X*, *Z* and *T* are incidence matrices for *h*, *a* and *p*, respectively.

The expectation and variance matrix of the random variables are defined as

$$E \begin{bmatrix} a \\ p \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \text{Var} \begin{bmatrix} a \\ p \\ e \end{bmatrix} = \begin{bmatrix} A\sigma_a^2 & 0 & 0 \\ 0 & I\sigma_p^2 & 0 \\ 0 & 0 & I\sigma_e^2 \end{bmatrix},$$

where σ_a^2 , σ_p^2 and σ_e^2 represent additive genetic variance, litter variance and residual variance, respectively, and *A* is the additive genetic relationship matrix for animals. The MME corresponding to this model is

$$\begin{bmatrix} X'X & X'Z & X'T \\ Z'X & Z'Z + A^{-1}k_1 & Z'T \\ T'X & T'Z & T'T + Ik_2 \end{bmatrix} \begin{bmatrix} \hat{h} \\ \hat{a} \\ \hat{p} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \\ T'y \end{bmatrix}, \tag{2}$$

Table 1 Number of records, herds and boars in three data sets

Breed	No. of herds	No. of records	No. of boars
Large White	36	31 459	2385
Landrace	27	12 982	1287
Duroc	18	8158	903

where $k_1 = \sigma_e^2/\sigma_a^2$ and $k_2 = \sigma_e^2/\sigma_p^2$. For analyzing the connectedness, only the coefficient matrix is needed.

Calculation of connectedness ratings

The CR between herd *i* and herd *j* is defined by the following formula (Mathur *et al.*, 1998):

$$CR_{ij} = \frac{\text{Cov}(\hat{h}_i, \hat{h}_j)}{\sqrt{\text{Var}(\hat{h}_i)\text{Var}(\hat{h}_j)}}, \tag{3}$$

where \hat{h}_i and \hat{h}_j are the estimated effects of herd *i* and herd *j*, respectively, obtained from the MME. The average CR for one herd is defined as the average of its CRs with all other herds.

The calculation of the variances and covariances in formula (3) requires the elements of the inverse of the coefficient matrix of MME corresponding to the herd effects. It is very difficult to obtain the inverse directly when the data set is large. Hence, Mathur *et al.* (1998) applied the following procedure:

Since

$$W'W(W'W)^{-1} = I$$

therefore

$$W'W(W'W)_i^{-1} = I_i, \tag{4}$$

where *W'W* is the coefficient matrix of MME, *I* is an identity matrix, $(W'W)_i$ is a vector of *W'W* corresponding to herd *i* and *I_i* is a vector of the identity matrix corresponding to herd *i*. Solving equations (4) for herd *i* and herd *j* separately, the required elements for calculation of the variances and covariances in formula (3) can be obtained. The program for calculating CR using this approach can be found at the website of the Canadian Center for Swine Improvement (<http://www.ccsi.ca/connectedness>).

Result and discussion

The average CR values for Large White, Landrace and Duroc are given in Tables 2–4, respectively. Only the herds with an average CR greater than zero are listed in Tables 2–4 as they are connected with at least one other herd.

In Large White, of the 36 herds analyzed there are 20 herds with average CR greater than zero. The highest CR (71.18%) was observed between herd BJHD1 and BJHD2. A further inspection into the data revealed that these herds shared 76 common sires that have 72.3% of their progeny in BJHD1 and 21.4% in BJHD2. Herd BJXD1 has the highest average CR because it shares its boars with four other herds. However, most of the other herds have strong genetic links with only one or few herds, but no or very weak connections with other

Table 2 Average connectedness rating between herds in Large White

Herd code	Region	Connectedness rating (%)		No. of herds connected	Most connected herd
		Maximum	Average		
BJXD1	BJ	59.78	3.59	8	BJZY
BJXD2	BJ	12.13	0.75	8	BJXD1
BJHD1	BJ	71.18	3.34	8	BJHD2
BJHD2	BJ	71.18	3.04	8	BJHD1
BBSC	BJ	27.9	0.82	7	BJCP
BJCP	BJ	27.9	0.90	7	BBSC
BJDX	BJ	6.16	0.39	7	BJXD1
BJTZ	BJ	23.80	1.45	8	BJXD1
BJZY	BJ	59.78	3.69	8	BJXD1
BSPF	GD	5.29	0.15	1	HYPF
HYPF	GD	5.29	0.15	1	BSPF
SCBF	SC	14.20	0.41	1	DYZC
DYZC	SC	14.20	0.41	1	SCBF
HBTZ	HB	1.20	0.04	3	HBMG
HBMG	HB	1.94	0.14	3	ZHZC
NHPF	TJ	13.51	0.40	2	QJPF
YNPF	YN	1.04	0.04	2	QJPF
TMPF	FJ	1.81	0.05	3	HBMG
QJPF	AH	13.51	0.42	2	NHPF
ZHZC	HN	1.94	0.06	3	HBMG

Table 3 Average connectedness rating between herds in Landrace

Herd code	Region	Connectedness rating (%)		Most connected herd	
		Maximum	Average	Herd code	Region
BBSC	BJ	42.22	1.62	BJCP	BJ
BJCP	BJ	42.22	1.62	BBSC	BJ
NHPF	TJ	3.96	0.15	QJPF	AH
QJPF	AH	3.96	0.15	NHPF	TJ

Table 4 Average connectedness rating between herds in Duroc

Herd code	Region	Connectedness rating (%)		Most connected herd	
		Maximum	Average	Herd code	Region
HAMG	HN	3.56	0.21	HAZW	HN
HAZW	HN	3.56	0.21	HAMG	HN
HBTZ	HB	1.27	0.07	QJPF	AH
SJYZ	HB	0.34	0.02	HBTZ	HB
QJPF	AH	1.27	0.07	HBTZ	HB

herds. In general, the higher connectedness was observed between herds within the same region.

In Landrace and Duroc, of the 27 and 18 herds analyzed, respectively, only four Landrace and five Duroc herds were found to have connectedness with other herds. They are all connected with just one herd.

A certain degree of connectedness between herds is a precondition for comparing EBVs of different herds with reasonable accuracy. However, as can be seen in the results, most major pig breeding herds in China have no or very weak connectedness with other herds. The major reason for

Table 5 Herds for which joint genetic evaluation can be conducted based on their connectedness rating

Group/ breed	Large White		Landrace		Duroc			
	Large White	Large White	Landrace	Landrace	Duroc	Duroc		
1	BJXD1	BJXD2	BBSC	BJCP	BBSC	BJCP	HAMG	HAZW
	BJHD1	BJHD2	BJZY	BJTZ				
2	BSPF	HYPF			NHPF	QJPF		
3	SCBF	DYZC						
4	NHPF	QJPF						

this is that there is very limited use of artificial insemination (AI) in the swine industry in China. The existing connectedness between the few herds is mainly due to the importation of breeding animals from a same foreign farm or sales of breeding animals from one farm to another. Figure 1 shows the flow of boars among herds that are connected to each other. As can be seen, the superior boars often came from same herds and the flow directions are unexceptionally unidirectional. There are some herds between which there is no direct link due to boars but still some connectedness exists, e.g., herds BSPF and HYPF. This may be due to sales of sows between them.

By analyzing the field data from the Canadian national swine database, Mathur *et al.* (1998) recommended that for backfat and age at 100 kg, a minimum average CR of 3% would be required to ensure a reasonably accurate comparison of EBVs of a herd with that of the other herds. According to this criterion, only for a few groups of herds as shown in Table 5, within which the average CR of one herd with other herds is greater than 3%, the EBVs of animals of one

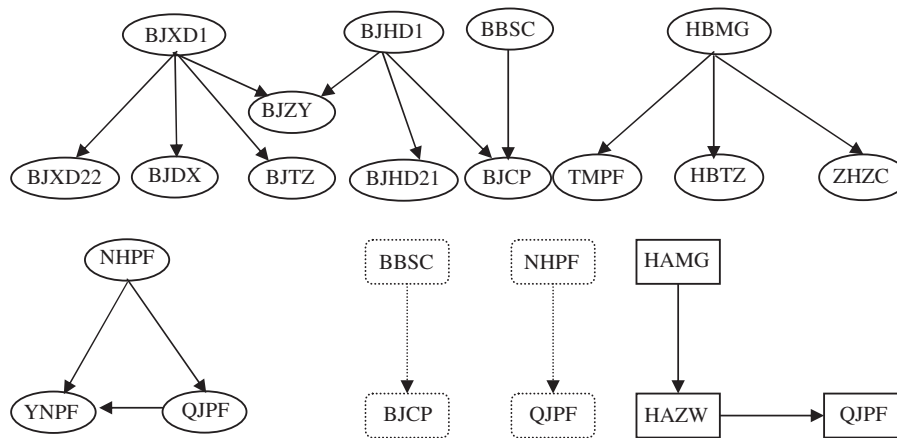


Figure 1 Flow of boars among herds that are connected to each other. Ellipses represent herds for Large White, rectangles for Duroc and rounded rectangles for Landrace.

herd are comparable with that of other herds of the same group with reasonable accuracy. This means that the cross-herd genetic evaluation has to be limited within these groups and a large-scale cross-herd genetic evaluation is not practicable at present. To implement the national joint genetic evaluation, great efforts are needed for improving the connectedness among herds in China. Schemes to increase the connectedness generally involve either the use of common sires across herds or the use of central test stations. In China, currently there are only a few test stations, but contrary to the situation in most countries of north America and Europe, where the numbers of test stations have declined, the number of test stations has increased in recent years and will be further increased in the near future. However, due to the limited scale of test stations, the role of test stations in increasing connectedness is limited. Therefore, the key point is to increase the use of common boars through extending the use of AI in the swine industry in China. Based on an extensive AI system, the connectedness can be improved by exchanging semen between herds or by forming a pool of superior AI boars, using AI boars with high CRs, and obtaining a sufficient proportion of progeny from these animals. Enough progeny of every reference sires is important to permit accurate comparisons of animals across herds. Increasing the proportion of progeny from the common sires can improve connectedness (Mathur *et al.*, 1998), and reduced the bias of comparisons (Kuehn *et al.*, 2007).

Conclusions

The levels of connectedness between pig breeding herds in China are generally low. Most herds are connected with only one or a few other herds. National or regional cross-herd genetic evaluation is not feasible currently. The cross-herd genetic evaluation can be conducted only for a few herds that have average CRs greater than 3% between them. Great efforts are needed to establish and enhance the connectedness between herds by means of extensively using AI in the swine industry in China. This will help to increase the accuracy

and effectiveness of across-herd genetic evaluations as well as to increase the rate of genetic improvement in China.

Acknowledgments

This study is financially supported by the 948 Project of the Ministry of Agriculture (Grant No. 2004-Q3) and the National Key Technology R&D Program of China (Grant No. 2006BAD10A02-01).

References

Foulley JL, Bouix J, Goffinet B and Elsen JM 1990. Connectedness in genetic evaluation. In *Advances in Statistical Methods for Genetic Improvement of Livestock* (ed. D Gianola and K Hammond), pp. 302–337. Springer Publishing, Heidelberg.

Foulley JL, Hanocq E and Boichard D 1992. A criterion for measuring the degree of connectedness in linear models of genetic evaluation. *Genetics Selection Evolution* 24, 315–330.

Fries LA 1998. Connectability in beef cattle genetic evaluation: the heuristic approach used in MILC.FOR. *Proceedings of the 6th World Congress of Genetics Applied to Livestock Production*, Communication No. 27449, Armidale, Australia.

Fu Y 2005. Swine genetic improvement program in China. *Proceedings of the National Swine Improvement Federation Conference*, Ottawa, Ontario, Canada.

Huisman AE, Tier B and Brown DJ 2006. On assessing contrasts between groups of animals. *Livestock Science* 104, 254–267.

Kennedy BW and Trus D 1993. Consideration on genetic connectedness between management units under an animal model. *Journal of Animal Science* 71, 2341–2352.

Kuehn LA, Lewis RM and Notter DR 2007. Managing the risk of comparing estimated breeding values across flocks or herds through connectedness: a review and application. *Genetics Selection Evolution* 39, 225–247.

Laloë D 1993. Precision and information in linear model of genetic evaluation. *Genetics Selection Evolution* 25, 557–576.

Laloë D, Phocas F and Ménéssier F 1996. Considerations on measures of precision and connectedness in mixed linear model of genetic evaluation. *Genetics Selection Evolution* 28, 359–378.

Mathur PK, Sullivan BP and Chesnais J 1998. Estimation of the degree of connectedness between herds or management groups in the Canadian swine population. *Canadian Centre for Swine Improvement*, Ottawa, Canada. (Mimeo).

Mathur PK, Sullivan BP and Chesnais JP 2002. Measuring connectedness: concept and application to a large industry breeding program. *Proceeding of the 7th World Congress of Genetics Applied to Livestock Production*, Communication No. 20-13, Montpellier, France.

Roso VM, Schenkel FS and Miller SP 2002. Estimation of the degree of connectedness between test groups of station-tested beef bulls. *Proceedings of the 7th World Congress on Genetics Applied to Livestock Production*, Communication No. 02-28, Montpellier, France.