

## Use of relationship matrix in the evaluation of natural service Limousin bulls

D Laloë, G Renand, J Sapa, F Méniéssier

*Institut National de la Recherche Agronomique, Station de Génétique  
Quantitative et Appliquée, Centre de Recherches de Jouy-en-Josas,  
78352 Jouy-en-Josas Cedex, France*

(Received 8 April 1991; accepted 12 December 1991)

**Summary** – In France, natural service bulls of beef breeds are progeny tested using reference sires and planned matings. The evaluation is made with a single-trait best linear unbiased prediction (BLUP) procedure applied to a sire model on pre-weaning calf performance, where sires are considered as unrelated. The purpose of this investigation was to examine the impact of using relationships between sires on breeding value estimates, their accuracies, and the connectedness between sires. Two mixed models, with or without sire interrelations were used. Weaning weights of 5 207 Limousin calves from 219 sires were analysed. A total of 53 sires were related by a Malécot relationship coefficient  $\geq 0.25$ . The correlation between breeding value estimates obtained with the 2 models was 0.969. The gain in accuracy was characterized by the reduction of prediction error variances and expressed by the ratio of these variances obtained with the 2 models. This gain averaged 4%, reached 26% and was related to the sire relationship coefficient means (correlation of 0.63). Overall connectedness indexes were computed and their interpretation discussed. The practical interest of using the relationship matrix in these evaluation programmes was examined in the light of French beef cattle breeding.

**relationship matrix / connectedness / accuracy / sire model / beef cattle**

**Résumé** – *Prise en compte des parentés pour l'évaluation des taureaux Limousins de monte naturelle. En France, les taureaux de races à viande de monte naturelle sont évalués sur descendance en utilisant des taureaux de référence et des accouplements planifiés. La méthode d'évaluation est un BLUP unicaractère appliqué à un modèle père sur les performances avant sevrage, où les taureaux sont considérés comme non apparentés. Cet article étudie l'incidence de la prise en compte des parentés entre taureaux sur l'estimation de leurs valeurs génétiques, la précision des estimées et la connexion entre taureaux. Deux modèles mixtes différant par la prise en compte ou non de la parenté sont considérés. Les poids au sevrage de 5 207 veaux Limousins, issus de 219 pères, sont inclus dans l'analyse. Cinquante-trois pères sont liés entre eux par un coefficient de parenté de Malécot supérieur ou égal à 0,25. La corrélation entre les valeurs génétiques obtenues dans les 2 modèles*

*est de 0,969. Le gain de précision, caractérisé par la réduction des variances d'erreur de prédiction, est exprimé par le rapport de ces variances obtenues dans les 2 modèles. Le gain se situe en moyenne à 4% et peut atteindre 26%. Il est lié à la moyenne des coefficients de parenté entre pères (corrélation de 0,63). Divers indicateurs du degré de connexion sont calculés et leur interprétation est discutée. L'intérêt pratique de considérer la parenté entre taureaux dans ces programmes d'évaluation est examiné à la lumière de la situation française.*

**matrice de parenté / connexion / modèle «père» / bovins viande / précision**

## INTRODUCTION

An efficient evaluation of natural service bulls in the French beef breeds is difficult to obtain using field performance, mainly because the number of contemporary sires within the herd, and the number of sires used in more than one herd is limited due to the small proportion of artificial insemination.

With such a data structure, disconnectedness results. Consequently, bulls used in different herds or years cannot be compared to each other, or comparison will be carried out imprecisely.

Therefore, a system for evaluating the breeding values of natural service beef sires was set up in 1980 in order to reduce this disconnectedness (Foulley and Sapa, 1982; Sapa and Ménissier, 1987; Ménissier, 1988). It is based on planned matings and artificial insemination with reference sires, using cows of different ages and morphological types within the same intra-herd period in order to obtain a random choice of mated cows. The evaluation concerns weight at birth and at 210 days as well as muscle and skeleton development scores at weaning. The method used is a single-trait BLUP applied to a sire model including the following effects: herd-year, sire, dam parity and morphology, sex-feed supplementation level, calf birth season. The sires are considered as unrelated. A selection index has been established and a connectedness index between all the sires is computed (Foulley *et al*, 1984) to obtain an automatic ranking of all sires exhibiting connectedness indexes above a given level (Foulley *et al*, 1984, 1990).

The increasing use of the best dams and sires of the herd or the breed leads to relationships among these sires. The use of relationships among sires is an alternative method to improve the beef sire evaluation (Henderson, 1975b; Slinger and Lewis, 1986). It was therefore decided to use these relationships in the evaluation. The purpose of this investigation was to examine the impact of relationships on the ranking of sire breeding values, their accuracies and connectedness between the sires in the French system.

## MATERIALS AND METHODS

### *Methods*

Unbalanced designs occur when observations in each level of one factor of the model are not equally distributed across levels of some other factor. In the extreme,

unbalancedness can result in disconnectedness, *ie*, contrasts between levels of one factor are no longer estimable. For random factors, the contrasts are always estimable, but the variance of the contrasts will increase with the unbalancedness of the design, which we called here "degree of disconnectedness".

Consequently, comparison of the degree of disconnectedness between BLUP estimates obtained with different models consists of comparing variances of contrasts between estimates: the greater these variances are, the greater the degree of disconnectedness from one model to another. The variances are quadratic forms associated with the variance-covariance matrices of prediction errors, and we will therefore examine quadratic form ratios. Two models were used involving the same effects, but differing in terms of the variance-covariance matrix of genetic effects (sire effects), *ie*, *model a* including between-sire relationships and *model i* not including these relationships. In order to study the impact of the relationship matrix, **A**, on the degree of disconnectedness, we used a third model, *model a'*, where all the off-diagonal coefficients of **A** were 0.25, and the diagonal coefficients were 1, corresponding to a population where the animals are non-inbred, and, for instance, half sibs.

### Models of analysis

The model is:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e}$$

where **y** is the vector of calf performance;

**b** is the vector of fixed effects (herd-year, sex-level of feed supplementation, calf birth season, dam parity) and of covariables (withers height and dam muscle development score, age of calves at scoring);

**u** is the vector of direct additive random effects of sires (variance  $\sigma_a^2$ );

**e** is the vector of random residuals (variance  $\sigma_e^2$ );

**X** and **Z** are incidence matrices relating effects to observations.

The first 2 moments of the random variables are:

$$E \begin{vmatrix} \mathbf{u} \\ \mathbf{e} \end{vmatrix} = \begin{vmatrix} \mathbf{0} \\ \mathbf{0} \end{vmatrix} \quad V \begin{vmatrix} \mathbf{u} \\ \mathbf{e} \end{vmatrix} = \begin{vmatrix} \mathbf{D}\sigma_a^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{I}\sigma_e^2 \end{vmatrix}$$

According to whether sire relationships are taken into account or not, **D** will be equal to the relationship matrix **A** (*model a*, *model a'*) or the identity matrix **I** (*model i*). We then assume that **e** is the same in the 3 models.

### Computation of **A** and **A**<sup>-1</sup>

**A** was computed according to the recursive procedure described by Henderson (1976). **A**<sup>-1</sup> was obtained by direct inversion of **A**.

For each sire, 2 relationship criteria have been computed, the mean relationship coefficient ( $r_m$ ) and the maximum relationship coefficient ( $r_M$ ).

### Mixed model equations

After absorption of the fixed effects, the mixed model equations are:

$$\begin{aligned} & (\mathbf{Z}'\mathbf{M}\mathbf{Z} + \lambda\mathbf{D}^{-1}) \hat{\mathbf{u}} = \mathbf{Z}'\mathbf{M}\mathbf{y} \\ \text{where} & \quad \mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \\ \text{and} & \quad \lambda = \sigma_e^2/\sigma_a^2 \end{aligned}$$

### Computation of variance-covariance matrices of prediction errors

Suppose that the true between-sires variance-covariance matrix is  $\mathbf{A}\sigma_a^2$ , the variance-covariance matrix of prediction errors ( $\mathbf{V}$ ) is obtained (Henderson, 1975a) according to the model by:

\* *model a, model a'*:

$$\begin{aligned} \mathbf{V}_a &= \mathbf{C}_a\sigma_e^2 \\ \text{where} & \quad \mathbf{C}_a = (\mathbf{Z}'\mathbf{M}\mathbf{Z} + \lambda\mathbf{A}^{-1})^{-1} \end{aligned}$$

\* *model i*:

$$\begin{aligned} \mathbf{V}_i &= \mathbf{C}_i\sigma_e^2 \\ \text{where} & \quad \mathbf{C}_i = \mathbf{C}_i^* + \lambda\mathbf{C}_i^*(\mathbf{A} - \mathbf{I})\mathbf{C}_i^* \\ \text{and} & \quad \mathbf{C}_i^* = (\mathbf{Z}'\mathbf{M}\mathbf{Z} + \lambda\mathbf{I})^{-1} \end{aligned}$$

Note that *model a* supplies BLUP estimates of sire effects, while *model i* does not. Hence, they are of minimum variance in the class of unbiased linear estimators. Consequently, the value of any quadratic form associated with  $\mathbf{C}_a$  will be less than or equal to the same quadratic form associated with the variance-covariance matrix of prediction errors resulting from another unbiased linear estimator.

In particular, this is true for the estimator provided by solving the mixed model equations corresponding to *model i*. Thus,  $\mathbf{L}'\mathbf{C}_a\mathbf{L} \leq \mathbf{L}'\mathbf{C}_i\mathbf{L}$  for any vector  $\mathbf{L}$ .

These matrices being both positive definite, the ratio  $\mathbf{L}'\mathbf{C}_i\mathbf{L}/\mathbf{L}'\mathbf{C}_a\mathbf{L}$  will be  $\geq 1$ .

### Evaluation of gain in accuracy of estimates

The variance of prediction error of the breeding value estimate of sire  $t$  is:

$$\text{var}(\hat{u}_t) = C_{(t,t)}\sigma_e^2,$$

and the gain in accuracy provided by considering relationships can be computed for sire  $t$  by the ratio:

$$C_{i(t,t)}/C_{a(t,t)}.$$

The correlations between this ratio and the 2 criteria related to the relationship coefficients of the sires ( $r_m$  and  $r_M$ ) were computed.

The overall gain in accuracy was the ratio between the traces of both matrices:

$$\text{tr}(\mathbf{V}_i)/\text{tr}(\mathbf{V}_a).$$

### Evaluation of the gain in connectedness between sires

The connectedness improvement and hence the gain in between-sire comparability was estimated by the ratio between quadratic forms:

$$\mathbf{L}'_t \mathbf{C}_i \mathbf{L}_t / \mathbf{L}'_t \mathbf{C}_a \mathbf{L}_t,$$

where  $\mathbf{L}_t$  is the contrast between the breeding value of sire  $t$  and that of all the other sires. This ratio varies from 1 to infinity and increases with the improvement of the connection between sire  $t$  and the other sires. The correlation between the connection improvement of 1 sire and its 2 relationship criteria ( $r_m$  and  $r_M$ ) can thus be computed as well as its gain in accuracy.

### Overall gain in connectedness

Furthermore, because  $\mathbf{C}_a$  and  $\mathbf{C}_i$  are positive definite matrices, the ratios of the quadratic forms associated with these 2 matrices ranged between  $\min(\mathbf{k})$  and  $\max(\mathbf{k})$ ,  $\mathbf{k}$  being solutions to the equation:

$$|\mathbf{C}_i - \mathbf{kC}_a| = 0$$

Thus, Foulley *et al* (1990) suggested 2 indicators of the degree of connectedness between sires depending on the following solutions:

$$\gamma_1 = \text{tr}(\mathbf{C}_a^{-1} \mathbf{C}_i) / n \quad \text{and} \quad \gamma_2 = \det(\mathbf{C}_i) / \det(\mathbf{C}_a)$$

where  $n$  is the number of sires (bulls),  $\gamma_1$  is the arithmetic mean of solutions  $\mathbf{k}$ , and  $\gamma_2$  represents the product of solutions  $\mathbf{k}$ .

Instead of  $\gamma_2$ , we suggest the geometric mean of solutions  $\mathbf{k}$  as an indicator of the degree of connection, *ie*:

$$\gamma_3 = (\gamma_2)^{1/n},$$

noting that  $\gamma_3$  is  $\leq \gamma_1$ , the difference between the 2 values reflecting the dispersion of solutions  $\mathbf{k}$ . The further these indexes are from 1, the larger the mean gain in connectedness resulting from the consideration of sire relationships.

### Animals

The data used in this study were those included in the 1990 evaluation program involving on-farm progeny testing of Limousin natural service sires. The entire data file was used. The evaluation concerned 219 bulls, *ie*: 8 reference sires and 211 natural service sires which had been progeny tested during the last 5 years. The number of evaluated natural service sires ranged from 1–4 per farm, the average

number being 1.8. A total of 5 207 progeny were tested involving an average 19 calves per natural service sire and 159 calves per reference sire. In the last 5 years, 227 herd-year effects were considered, each involving an average of 6 progeny from reference sires and 17 from natural service sires.

Because of the objective of this study, only calf weight at 210 days was analysed. The heritability value used for this trait was 0.20, which corresponds to previous estimates (Laloë *et al*, 1988). The pedigrees of the bulls were established from the Limousin breed file, provided by "UPRA-France-Limousin-Sélection".

## RESULTS AND DISCUSSION

### *Relationship matrix structure*

The relationship matrix between the 219 sires was computed from their known ancestry over 5 generations, *ie*, a total of 2 144 animals. The structure of **A**, the matrix of relationship between these 219 sires is given in table I. The coefficients in **A** are twice the Malécot coefficients.

**Table I.** Distribution of relationship matrix **A** coefficients among 219 sires.

Value		Number	Frequency (%)
Off diagonal:			
	0	15 354	64.32
]0	to 0.062 5[	7 862	32.94
[0.062 5	to 0.125[	376	1.58
[0.125	to 0.25[	170	0.71
[0.25	to 0.50[	76	0.32
[0.50	to 1[	33	0.14
Total		23 871	100.00
Diagonal:			
	1	160	73.06
	> 1	59	26.94
Total		219	100.00

Almost 2/3 of the relationship coefficients were null and only 3.6% of the pairs of sires exhibited a relationship coefficient  $\geq (1/2)^4$ . The 33 relationship coefficients  $\geq 0.5$  belonged to 53 sires (sire-son pairs). In 1986 (H Roy, personal communication), a similar relationship structure was already observed between Limousin sires of this evaluation programme.

### **Breeding values**

Elementary statistics concerning the breeding values are given in table II. Their distribution did not markedly differ whether sire relationships were used or not.

The correlation coefficients between the breeding values of the 2 evaluations were 0.969 (Pearson's coefficient) and 0.961 (Spearman's correlation). The correlation coefficients computed from the 53 sires related to each other by a relationship

**Table II.** Statistics of breeding value estimates of sires according to evaluation models.

<i>Model</i>	<i>Mean</i>	<i>SD</i>	<i>Minimum</i>	<i>Maximum</i>
Unrelated ( <i>model i</i> )	0	3.37	-11.71	10.68
Related ( <i>model a</i> )	0.03	3.45	-12.20	9.90

Unit: 210 day weight (kg). SD: standard deviation.

coefficient  $\geq 0.5$  were 0.920 (Pearson's correlation) and 0.907 (Spearman's correlation). Although these coefficients indicate a close relationship between the 2 evaluations, the individual breeding values nevertheless exhibited substantial differences, as shown in table III. Means and maxima of the absolute value of the differences both in terms of sire breeding values and rankings increased with the mean relationship level (mean of sire relationship coefficients). The ranking of sires with extreme breeding values remained similar.

**Table III.** Mean differences and gain in accuracy and connectedness according to mean relationships.

<i>No of sires:</i>	<i>Mean relationship coefficient: (<math>r_m</math>)</i>	<i>Absolute difference: means</i>		<i>Mean gain of <sup>a</sup>:</i>	
		<i>Breeding values (kg)</i>	<i>Rank order</i>	<i>Accuracy</i>	<i>Connectedness</i>
97	[0.009 to 0.015[	0.197	4.91	1.015	1.015
74	[0.015 to 0.020[	0.741	14.86	1.044	1.043
38	[0.020 to 0.025[	0.920	16.97	1.066	1.063
10	[0.025 to 0.030[	1.258	22.10	1.143	1.136
Average	0.016	0.555	11.15	1.038	1.040
Maximum		(3.329) <sup>b</sup>	(70) <sup>b</sup>	1.255	1.253

<sup>a</sup> Criteria defined in the paper (variance ratio: model without relationships/model with relationships); <sup>b</sup> Maximum deviation.

### Gain in accuracy

The average gain in accuracy per sire was 1.038, ranging from 1 to 1.255 (table III). This gain in accuracy was correlated with the mean level of relationship ( $r = 0.63$ ) and with the maximum relationship ( $r = 0.81$ ) of the sire. As expected, the gain in accuracy increased with increasing mean relationship between sires. Comparing *model i* and *model a'*, the average gain in accuracy was 1.101, ranging from 1.082 to 1.252: the average gain in accuracy seems to increase with A.

### **Gain in connectedness**

The gain in connectedness per sire reached a mean value of 1.040 and ranged from 1 to 1.253 (table III). The gain increased with relationship coefficients, as shown by the correlation between this gain and  $r_m$  ( $r = 0.60$ ), and  $r_M$  ( $r = 0.81$ ).

There was a very high correlation ( $r = 0.998$ ) between gain in connectedness and gain in accuracy. However, comparing *model i* and *model a'*, the gain in connectedness per sire reached a mean value of 1.019 and ranged from 1 to 1.023. The correlation between gain in connectedness and gain in accuracy was only  $-0.22$ .

The values of indicators  $\gamma_1$  and  $\gamma_3$  of the overall degree of connectedness were 1.038 and 1.035, respectively. Comparing *model i* and *model a'*, the values of indicators  $\gamma_1$  and  $\gamma_3$  of the overall degree of connectedness were the same: 1.020.

### **CONCLUSION**

Use of the relationship matrix in the evaluation of natural service sires leads to marked estimation changes and also to an improvement (+4%) in the accuracy and connectedness of estimates. An increasing relationship matrix increased the precision of the evaluations, but not the degree of connectedness. More studies are required to study the impact of including relationships on the degree of connectedness on genetic evaluations.

### **ACKNOWLEDGMENTS**

We are very grateful to UPRA-France-Limousin-Sélection who supplied pedigree data and to Mrs K Rérat (INRA-UCD, Jouy-en-Josas) for the translation of the manuscript into English.

### **REFERENCES**

- Foulley JL, Sapa J (1982) The French evaluation program for natural service beef bulls using AI sire progeny as herd ties. *Br Cattle Breed Club Winter Conf, Jan 1982*. Cambridge, UK, 10 pp
- Foulley JL, Bouix J, Goffinet B, Elsen JM (1984) Comparaison des pères et connexion. *In: Insémination Artificielle et Amélioration Génétique: Bilans et Perspectives Critiques* (Elsen JM, Foulley JL, eds) Les Colloques de l'INRA, INRA Publ, Paris, 29, 133-176
- Foulley JL, Bouix J, Goffinet B, Elsen JM (1990) Connectedness in genetic evaluation. *In: Advances in Statistical Methods for Genetic Improvement of Livestock* (Gianola D, Hammond K, eds) Springer, Heidelberg, 302-337
- Henderson CR (1975a) Comparison of alternative sire evaluation methods. *J Anim Sci* 41, 760-770
- Henderson CR (1975b) Use of relationships among sires to increase accuracy of sire evaluation. *J Dairy Sci* 58, 1727-1732
- Henderson CR (1976) A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* 32, 60-83



Laloë D, Renand G, Ménissier F, Astruc JM, Gaillard J, Sapa J (1988) Genetic parameters among preweaning calf performance for beef breeds in French field record programs. In: *Proc 3rd World Congr Sheep and Beef Cattle, June 1988, Paris, France*. INRA Publ, Paris, vol 1, 355-358

Ménissier F (1988) La sélection des races bovines à viande spécialisées en France. *Proc 3rd World Congr Sheep and Beef Cattle, June 1988, Paris, France*. INRA Publ, Paris, vol 2, 215-236

Sapa J, Ménissier F (1987) L'évaluation et la sélection en France des taureaux de races à viande destinés à la monte naturelle. II. Le contrôle sur descendance en ferme. *Bull Tech CRZV Theix* 69, 21-29

Slanger WD, Lewis BG (1986) Including pedigree relationship in beef sire evaluation. *J Anim Sci* 62, 1563-1575