

## COMPARISON OF ALTERNATIVE MODELS FOR SELECTION FOR LITTER SIZE IN DANISH LANDRACE AND YORKSHIRE BREEDS

J. Estany<sup>1</sup> and D. Sorensen

National Institute of Animal Science, Foulum, Postbox 39, 8830 Tjele, Denmark

<sup>2</sup> Centre UdL-IRTA, Rovira Roure 177, 25006 Lleida, Spain

### SUMMARY

The model to predict breeding values for selection for litter size in the Danish pig breeding program was reviewed. A repeatability animal model with genetic groups was compared to a model without groups, being the groups defined according to the country of importation. Within each case, the herd by year by type of insemination effects (H) were fitted as fixed or random. It is concluded that predictive ability for litter size was improved by treating H effects as random, but there was no evidence that genetic groups contribute to enhance it.

### INTRODUCTION

Since 1988, a repeatability animal model to predict breeding values for litter size has been available to Landrace and Yorkshire Danish pig breeders (Sorensen, 1991). The predictions are computed from data supplied by pig breeders from registered herds, and the model includes such effects as herd-year, season, parity number and type of conception. As importation of animals from different countries has taken place over years, the model also includes a genetic group effect which takes into account the fact that breeding values are sampled from populations whose mean could differ. The purpose of this paper is to compare the predictive ability of the model currently in use against other three related models. Thus four models for total number of born piglets are fitted to the data. Estimates of variances and related parameters are obtained under models with and without genetic groups, and treating herd-years as fixed or random.

### MATERIALS AND METHODS

The data available for this study were obtained from the Danish pig breeding program. The data set consisted of 19666 and 29336 litters from purebred Danish Landrace and Yorkshire respectively. These data sets are referred to as L and Y (Landrace, Yorkshire). The trait analyzed was total number of piglets born, as it is the trait currently evaluated in the national program. Individuals imported from an identified country were grouped according to the country of origin. Other individuals were assigned to the Danish group. In both breeds five genetic groups were included in the model. Total number of piglets born was analyzed using the following univariate repeatability model:

$$y_{ijkmno} = H_i + S_j + K_k + \sum x_m g_m + a_{mn} + p_{mn} + e_{ijkmno}$$

where  $y_{ijkmno}$  is the  $o$ th record (total number of piglets born) of the  $n$ th animal;  $H_i$  is the herd by year by type of insemination (natural vs artificial) effect  $i$  (from now onwards referred to as the herd-year effect);  $S_j$  is the effect of season  $j$  (four quarters within a

year);  $K_k$  is the effect of parity class  $k$  (first parity number was subdivided into 5 classes according to age at first parity);  $x_m$  is the contribution of the  $m$ th genetic group ( $\sum x_m = 1$ );  $g_m$  is the effect of the  $m$ th genetic group, defined as the country of origin;  $a_{mn}$  is the additive genetic value of animal  $m$  within group composition; and  $p_{mn}$  and  $e_{ijklmno}$  is the permanent environmental effect of the  $n$ th animal, and the error term associated with the  $o$ th record, respectively. Season, parity and group effects were treated as fixed whereas additive genetic, permanent environmental and residuals were treated as random. Herd-year effects were treated either as fixed or random, depending on the model used. Assumptions about group effects were described by Quaas (1988). Estimates of variance components obtained by restricted maximum likelihood (REML) were used in lieu of the true parameters. Variance components were estimated using the DFREML derivative-free approach (Meyer, 1988). Approximate F-tests were carried out, assuming that these estimated variances were the true parameter values.

Four different analyses were carried out within breed. In the first two, genetic groups were included in the model; in the other two, genetic groups were removed from the model. Within each case, herd-year effects were fitted as fixed or random. The four models will be referred to as follows: FG: herd-year fixed and with groups; RG: herd-year random and with groups; F: herd-year fixed and without groups; and R: herd-year random and without groups. Within each breed, the above data set (L and Y) was subdivided into two parts, whereby all the parameters of the model (including the variance components) were estimated using the first part, and then predicted records from the second part were computed. Using the observed and predicted records from the second part of the data set, the correlation between observed and predicted records ( $r_{yo,yp}$ ) and the average squared difference between observed and predicted records (MSE, mean squared error) were computed. These two criteria were used to compare the models. This is a test on the predictive ability of the models. Similar criteria were used by Pérez-Enciso and Gianola (1993). In order to study the predictive performance of different models with different amount of information available to estimate parameters, three arbitrary ways of subdividing the data were carried out. Thus in Landrace (Yorkshire), the L-1/2 (Y-1/2) subdivision resulted in two halves of equal size. The subdivision was carried out within litters, and only sows with at least two parities were included. In the L-1/5 (Y-1/5) subdivision and in the L-3/4 (Y-3/4) subdivision, parameters were estimated using either 20% and 75% of the (L and Y) data sets, and only 20% of the records were predicted.

## RESULTS AND DISCUSSION

Estimates of heritability with a univariate model, for Landrace, ranged from 11% to 14%, depending on the model. Excluding groups from the model resulted in biased estimators of heritability. As expected, the bias was positive and of the order of 13% in models where herd-year effects were treated as fixed, and 6% when they were treated as random. Repeatability estimates were about 19% for all models. For Yorkshire, heritability and repeatability was approximately 11% and 19% respectively, regardless of which model was fitted. Herd-year accounted for 2.6% (Landrace) and

3.9% (Yorkshire) of the total phenotypic variance.

Estimates of genetic group differences for number of piglets born are given in Table 1. In both breeds, estimates of differences between genetic groups were significant, according to the approximate F-test performed ( $P < 0.05$ ). Some of the estimates exceed 2 piglets. This value seems rather large and perhaps these estimates should be interpreted with reservation. Most of the imported animals consist of boars, and there are hardly any sows with records with more than 50% of "foreign" genes. If groups do not combine additively, as one would expect if animals originate from populations with different selection histories for litter size, estimates of groups differences are biased. No attempt was made in the present work to account for non-additive gene action.

In Table 2 are given the results for predictive ability in terms of  $r_{y_o, y_p}$  and MSE. The figures for  $r_{y_o, y_p}$  ranged from .31 to .39, increasing with the amount of information used to estimate parameters. The figures for MSE ranged from 7.2 to 9.3, being larger in Yorkshire, as it would be expected attending to the larger variance of this breed. The R and RG models improved MSE from 2% to 8%, the differences being higher in small than in big data partitions (7 vs 20 records per subclass). These figures show quite consistently that in both breeds models treating herd-year effects as random have better predictive ability than models treating them as fixed, measured both by  $r_{y_o, y_p}$  and by MSE, and that group effects have little influence on the predictive power of the model. This is an important criterion in the context of selection of animals as parents of the next generation, or selection of sows for future parities. It seems therefore justified to conclude that RG or R models should be used in the breeding program in the future.

Traditionally, herd-year effects in animal breeding are treated as fixed. A reason for this may be the argument put forward by Henderson (1973), who showed that predictors of breeding values, derived ignoring selection, can be biased, if there is an association between breeding values and random herd-year subclasses. The bias vanishes, if the random herd-year effects are treated as fixed. Henderson (1973,1975) based his conclusion using a repeated sampling approach, where he conditioned on the data structure and on the type of selection practiced. Jansen (1987) presented simulation results which challenged some of Henderson's (1975) conclusions. An alternative approach is to write down the likelihood or the posterior distribution, with and without selection, and to study under which set of conditions, they are mathematically identical. Using this approach, Gianola and Fernando (1986) and Im et al (1989) showed in an animal breeding context that selection is ignorable when it is based on the observed data. This implies that there is no need to treat herd-year effects as fixed, in order to draw correct inferences, if they are truly random, provided that information on which selection is based is included in the data.

#### REFERENCES

- GIANOLA, D. and FERNANDO, R.L. (1986) *J. Anim. Sci.* 63: 217-244  
HENDERSON, C.R. (1973) *Proc. Anim. Breed. Genet. Symp. in Honor Dr J.L. Lush*, pp. 10-41.  
HENDERSON, C.R. (1975) *Biometrics* 31: 423-447.

- IM,S.,FERNANDO,R.L. and GIANOLA,D. (1989) Genet. Sel. Evol 21:399-414  
 JANSEN, G.B. (1987) 38th Ann. Meet. EAAP, Lisbon.  
 MEYER,K. (1988) DFREML. User notes. University of Edinburgh.  
 PEREZ-ENCISO,M. and GIANOLA,D. (1993) Liv. Prod. Sci. 35:303-316.  
 QUAAS,R.L.(1988) J. Dairy Sci. 71:1338-1345.  
 SORENSEN,D.(1991) Report No 697, National Institute of Animal Science. Denmark.

**Table 1** Genetic groups estimates for total number of piglets born in Danish Landrace and Yorkshire, expressed as deviations from the Danish group<sup>(1)</sup>.

Data	Model	Genetic group			
		1	2	3	4
L	FG	0.73	2.08	1.37	-1.08
	RG	-0.29	1.69	1.04	-1.68
Y	FG	-1.90	1.45	0.39	-0.32
	RG	-1.85	1.38	0.63	-0.18

<sup>(1)</sup>F(R): model with herd-year as fixed (random)

**Table 2** Predictive analyses for total number of piglets born in terms of correlation between observed and predicted values ( $r_{yo,yp}$ ) and mean square error (MSE) in Danish Landrace and Yorkshire data partitions<sup>(1)</sup>.

Data set		FG	F	RG	R
L-3/4	$r_{yo,yp}$	0.349	0.351	0.389	0.390
	MSE	7.607	7.591	7.336	7.326
L-1/2	$r_{yo,yp}$	0.341	0.338	0.396	0.394
	MSE	7.783	7.804	7.240	7.253
L-1/5	$r_{yo,yp}$	0.316	0.317	0.365	0.367
	MSE	8.115	8.093	7.483	7.468
Y-3/4	$r_{yo,yp}$	0.315	0.315	0.343	0.343
	MSE	8.954	8.952	8.753	8.754
Y-1/2	$r_{yo,yp}$	0.319	0.318	0.365	0.364
	MSE	9.004	9.003	8.493	8.498
Y-1/5	$r_{yo,yp}$	0.306	0.305	0.345	0.344
	MSE	9.325	9.333	8.730	8.736

<sup>(1)</sup>F(R):model with herd-year as fixed (random); G:model with groups

**Acknowledgments.**

The first author gratefully acknowledges a research grant from DGICYT (Spain). K.Meyer and K.Boldman kindly provided copies of their computer programs.