Acta Veterinaria Hungarica 50 (4), pp. 373–383 (2002)

ANALYSIS OF THE RANDOM DISTRIBUTION OF STATION-TESTED PIGS BASED ON THEIR GENETIC MERIT

I. NAGY[1*], L. CSATÓ[1], J. FARKAS[2], L. RADNÓCZI[3], A. SZABÓ[4] and Zsófia VÍGH[1]

[1]Department of Pig Breeding, [2]Department of Mathematics and Informatics, [4]Diagnostic and Oncoradiologic Institute, University of Kaposvár, H-7400 Kaposvár, Guba S. u. 40, Hungary; [3]National Institute for Agricultural Quality Control, Budapest, Hungary

(Received January 10, 2002; accepted May 27, 2002)

The analysis was conducted on the national database of the station tests carried out between May 1996 and February 2001, using the Hungarian Large White (LW) and Hungarian Landrace (LR) breeds, and the A-line of the Ka-Hyb hybrid (MLW). Days of test, total amount of feed consumed and valuable cuts were taken into the analysis to test the random distribution of the animals' phenotypic measurements and predicted breeding values across the contemporary groups of origin. The phenotypic measurements of all traits were clearly positively associated with the herds of origin ($b = 0.52 - 1.08$). Animals with favourable phenotypic measurements originated from better herds. On the contrary, the predicted breeding values seemed to be independent of the herd effects ($b = -0.16 - 0.08$) and suggested a possible random distribution across the contemporary groups.

Key words: Pig, breeding value, bias, random distribution (across the herds)

By carrying out any breeding programme, one has to realise selection decisions based on measurements of certain traits (selection criteria) in order to achieve improvement in the traits of the breeding objective. Since the late 1980s several software packages have become available, which are capable of separating the phenotypic measurements into breeding values (additive genetic effects) and environmental effects even using extremely large data sets. These computer packages are based on the BLUP method, which is an efficient tool of selection and has been applied in various countries in breeding programmes of numerable species.

Unfortunately in the Hungarian pig breeding sector the BLUP procedure is only being used as an unofficial supplementary tool, helping selection decisions. The process of selection is still being accomplished using the station test index score (OMMI, 2000), which method is based on the Hazel index. That procedure, however, was developed almost sixty years ago (Hazel, 1943) and it is much less efficient than the BLUP method.

---

[*] Corresponding author: Dr. István Nagy; E-mail: nagy@mail.atk.u-kaposvar.hu; Fax: +36 (82) 320 175

The main obstacle hindering the exclusive application of the BLUP method regarding selection decisions is the strong resistance among the breeders against this procedure. They claim that the method unfairly overcompensates the environmental effects especially those of the herds (of origin). Thus, in their opinion, providing good conditions for the test animals before the station test takes place will be so to say 'punished' by the BLUP method and the resulting breeding values biased downwards. Accordingly, they claim the opposite to be true as well, that is providing bad conditions prior to the station test will be exaggeratedly recompensated causing breeding values being biased upwards.

Apart from personal beliefs BLUP can in fact be biased in case there is a nonrandom distribution of animals across the herds according to the animals' genetic merit (predicted breeding values) (Hofer, 1995). This situation can be attained in different ways. Either all the good animals (i.e. those with favourable breeding values) would originate in herds providing the best environmental conditions or animals with the highest genetic merit would be found in herds with the worst conditions prior to the station test (as claimed by the breeders).

Nevertheless, breeding value prediction is considered to be ideal when the distribution of animals is random across the herds according to the predicted breeding values, which simply means that animals both with high and low breeding values can be found having originated from both so-called good and bad herds, respectively.

With the present study the objective of the authors was to investigate the distribution of animals across the herds in relation to their genetic merit. Thus it may be determined whether or not the breeders' concern regarding the fairness of the BLUP method can be justified.

Materials and methods

Data recording

The genetic analysis was conducted on the data collected by the National Institute for Agricultural Quality Control of Hungary between May 1996 and February 2001, in the course of station tests. The Hungarian Large White (LW) and Hungarian Landrace (LR) breeds, and one synthetic Large White type line (the A-line) of the Ka-Hyb hybrid pig breeding programme (MLW) were analysed.

Station test (progeny test)

For the purposes of the station test a castrate and a female from the same litter are sent to the station between the age of 65 and 77 days. Body weight of the animals at the age of 65 days should be at least 17 kg but not higher than 32 kg. After a certain preliminary adaptation period the test begins at the age of 80 days (body weight at this age is at least 23 kg) and ends when reaching the final body weight of 105 kg. Body weight is measured at the beginning and at the end of the test with an accuracy of 1 kg. Animals are fed *ad libitum* and penned individually (OMMI, 2000). The traits chosen for analysis were days of test (DOT), total amount of feed consumed during the test (FEED) and valuable cuts (VC) (neck, shoulder, longissimus dorsi and ham), respectively. Basic statistics of the station test data are presented in Table 1.

Table 1

Basic statistics for the examined traits

Traits	Breed	No. of records	Mean	σ
DOT (days)[a]	LW[d]	6855	80.1	14.1
	LR[e]	2791	76.8	13.7
	MLW[f]	786	79.3	13.4
FEED (kg)[b]	LW	6855	202.1	22.2
	LR	2791	198.7	24.5
	MLW	786	204.3	21.3
VC (kg)[c]	LW	6855	39.2	2.59
	LR	2791	38.6	2.51
	MLW	786	40.3	2.75

[a]DOT, days of (station) test; [b]FEED, consumed feed; [c]VC, valuable cuts; [d]LW, Hungarian Large White; [e]LR, Hungarian Landrace; [f]MLW, one synthetic Large White type line (the A-line) of the Ka-Hyb hybrid pig breeding programme

Statistical analysis

The statistical analysis consisted of four consecutive steps. The first step was testing for the significance of the various environmental factors (fixed effects) conducting least squares analyses using the GLM procedure of the BMDP package leaving only significant factors in the model. Regarding the fixed effects, in the station test traits the tested fixed effects were herd (of origin), sex, and year-month (of the station test) and station.

The second step was the estimation of the variance and covariance components (random effects) of the examined traits. The method used to obtain the (co)variance components was the appropriate variation of the animal model (multitrait animal model) using the PEST (for data coding) (Groeneveld, 1990) and VCE 4 (Groeneveld, 1998) software packages (under LINUX) based on the REML method.

The residual, additive genetic (animal) and litter variances of DOT, FEED and VC were obtained using the following linear model:

$$y = Xb + Za + Wc + e$$

where (according to Mrode, 1996): y = vector of observations, b = vector of fixed effects, a = vector of random animal effects, c = vector of random litter effects, e = vector of random residual effects, X, Z and W are incidence matrices relating records to fixed effects, random animal and random litter effects, respectively.

Expected values of a, c and e were $E(a) = E(c) = E(e) = 0$.

The variance-covariance structure assumed to be $V(a) = A\sigma^2a$, $V(c) = I\sigma^2c$, $V(e) = I\sigma^2e$, and $cov(a,e) = cov(e,a) = cov(c,e) = cov(e,c) = 0$, where A is the numerator relationship matrix. Also $cov(y,a) = ZAI\sigma^2a$.

Distribution of y was assumed to be normal. Each trait was determined by many additive genes of infinitesimal effects at infinitely many unlinked loci. Residual, additive genetic and litter covariances were also estimated among the traits of DOT, FEED and VC respectively. Structure of the station test data is presented in Table 2.

Table 2

Structure of station test data

Breed	Herd	Sex	Year-month	Station	Total pedigree
LW[a]	57	2	59	7	12,521
LR[b]	34	2	56	7	5,385
MLW[c]	9	2	54	4	1,391

[a]LW, Hungarian Large White; [b]LR, Hungarian Landrace; [c]MLW, one synthetic Large White type line (the A-line) of the Ka-Hyb hybrid pig breeding programme

The third step was the breeding value prediction of the examined traits using the variance-covariance components obtained in the second step. Breeding value prediction was accomplished applying the PEST software (Groeneveld, 1990). The predicted breeding values [STOP = 0.0001, MAX_ITER = 3000] were considered as the measure of the genetic merit. At the same time the effects of the environmental factors of the examined traits were also estimated (BLUE). The estimated effects of the herds of origin served as the ranking figures between the herds.

Finally correlation coefficients were estimated between the breeding values (of all animals having phenotypic observations for the various traits) and the appropriate herd effects (from which the animals were originated) using the SPSS software (SPSS Inc., 1999). The lack of significance concerning the received correlation coefficients would provide the possibility for the existence of the random distribution of the test animals across the herds in relation to their genetic merit (i.e. in case the correlation coefficient is not significantly different from zero then the distribution may be random). Moreover, breeding values linearly regressed on the contemporary group effects in order to determine of the latter's possible effect on the former.

Results and discussion

Influencing factors

The tested fixed effects, namely herd (of origin), sex, and year-month (of the station test) and station effects all significantly influenced each of the examined traits (DOT, FEED, VC).

Variance-covariance estimates

The results of the second step were not the main interest of the present study, therefore it was decided that its results would not be published. However it has to be noted that using the same models and software packages (PEST, VCE) similar (co)variance components were received by Groeneveld et al. (1996) to those of obtained in the second step.

Correlation and linear regression coefficients

In order to improve perspicuity, apart from providing the correlation coefficients between the breeding values and contemporary group effects and linear regression coefficients of the predicted breeding values on the environmental effects of the herds of origin (Tables 3–4) the predicted breeding values (ordinate) and the appropriate herd effects (abscissa) were scatter plotted on Cartesian coordinates. All these procedures (correlation, regression, scatter plot) were repeated using the phenotypic measurements as dependent variables to create a basis for comparison (Figs 1–9).

Looking at the results in more detail, the first trait taken into the analysis was DOT (Figs 1–3, Tables 3–4). Based on the description of the station test (see above) the aim is clearly to complete the station test as quickly as possible reaching the target weight of 105 kg. Therefore comparing any values to each other either on the dependent or on the independent scale the smaller value is more advantageous. Viewing the results of the phenotypic measurements, a moderate positive association was found between the DOT measurements and the appropriate contemporary group effects. The positive correlation indicates a tendency of DOT measurements and herd effects to increase or decrease together. Accordingly a clear dependence was found (see on the left-side plots: Figs 1–3) concerning the duration of the test on the herds of origin effect. The linear regression coefficients were significantly different from zero (Tables 3–4). This phenomenon was very unfortunate because it means that the station test (on which the selection decision is based) is biased and gives an unfair advantage to those animals which are treated better than others prior to the test. However, this unfair advantage was corrected by the BLUP method and can also be seen on the right-side plots (Figs 1–3). The correlation and regression coefficients (Tables 3–4) were practically zero; thus, the predicted breeding values of the DOT were possibly not dependent on the environmental effects of the herds regardless of the genotype.

Table 3

Estimated correlation coefficients between the phenotypic measurements and herd effects and linear regression coefficients of the phenotypic measurements on the herd effects of the analysed pig populations' station test traits (level of significance of the estimates are given in brackets)

Traits	Breed	r[g]	b[h]
DOT (days)[a]	LW[d]	0.289***	0.882***
	LR[e]	0.350***	1.076***
	MLW[f]	0.208***	0.517***
FEED (kg)[b]	LW	0.381***	0.932***
	LR	0.399***	0.905***
	MLW	0.274***	0.679***
VC (kg)[c]	LW	0.190***	0.799***
	LR	0.249***	0.951***
	MLW	0.268***	0.653***

[a]DOT, days of (station) test; [b]FEED, consumed feed; [c]VC, valuable cuts; [d]LW, Hungarian Large White; [e]LR, Hungarian Landrace; [f]MLW, one synthetic Large White type line (the A-line) of the Ka-Hyb hybrid pig breeding programme; [g]correlation coefficient; [h]linear regression coefficient; ***estimate is significant at the 0.001 level

Table 4

Estimated correlation coefficients between the predicted breeding values and herd effects and linear regression coefficients of the predicted breeding values on the herd effects of the analysed pig populations' station test traits (level of significance of the estimates are given in brackets)

Traits	Breed	r[g]	b[h]
DOT (days)[a]	LW[d]	0.054***	0.036***
	LR[e]	–0.023 (NS)	–0.014 (NS)
	MLW[f]	0.007 (NS)	0.003 (NS)
FEED (kg)[b]	LW	0.000 (NS)	0.000 (NS)
	LR	0.023 (NS)	0.010 (NS)
	MLW	0.068 (NS)	0.031 (NS)
VC (kg)[c]	LW	–0.012 (NS)	–0.023 (NS)
	LR	0.043*	0.079*
	MLW	–0.130***	–0.159***

[a]DOT, days of (station) test; [b]FEED, consumed feed; [c]VC, valuable cuts; [d]LW, Hungarian Large White; [e]LR, Hungarian Landrace; [f]MLW, one synthetic Large White type line (the A-line) of the Ka-Hyb hybrid pig breeding programme; [g]correlation coefficient; [h]linear regression coefficient; ***estimate is significant at the 0.001 level; *estimate is significant at the 0.05 level; NS, nonsignificant

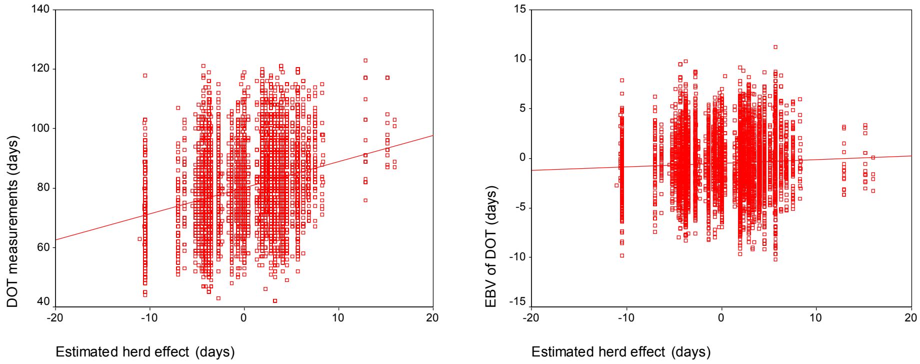

Fig. 1. Scatter plot of the Hungarian Large White test animals' DOT measurements and predicted breeding values (EBV) according to the test animals' herds of origin

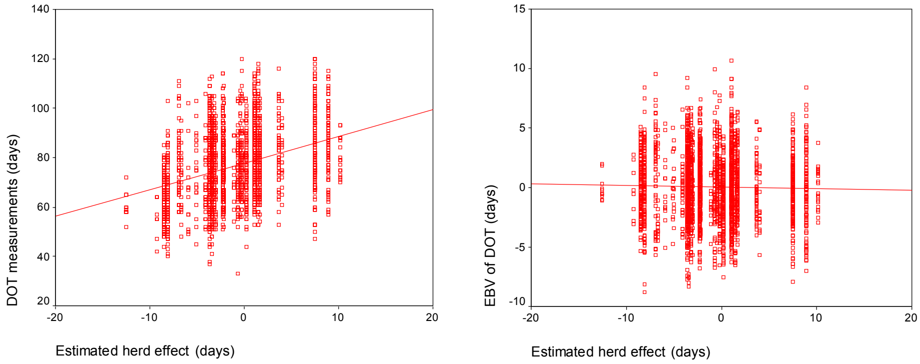

Fig. 2. Scatter plot of the Hungarian Landrace test animals' DOT measurements and predicted breeding values (EBV) according to the test animals' herds of origin

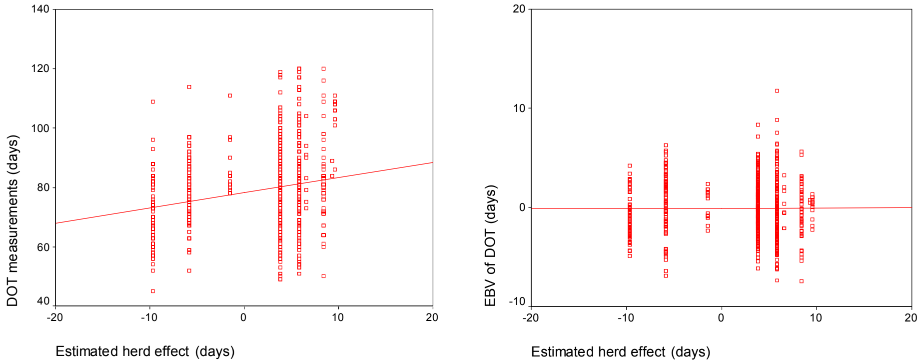

Fig. 3. Scatter plot of the Ka-Hyb hybrid (A-line) test animals' DOT measurements and predicted breeding values (EBV) according to the test animals' herds of origin

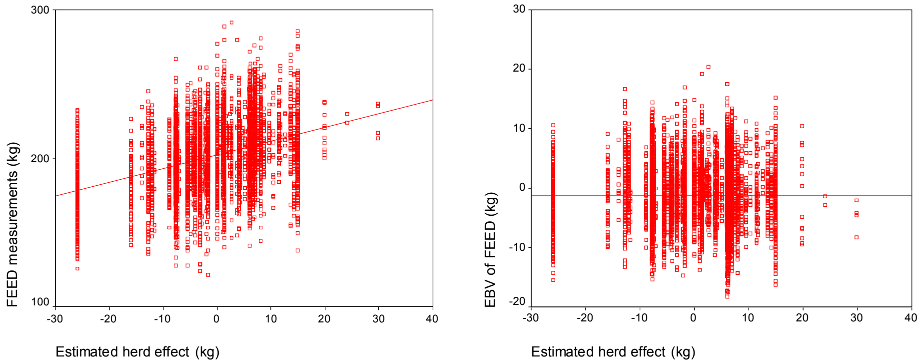

Fig. 4. Scatter plot of the Hungarian Large White test animals’ FEED measurements and predicted breeding values (EBV) according to the test animals’ herds of origin

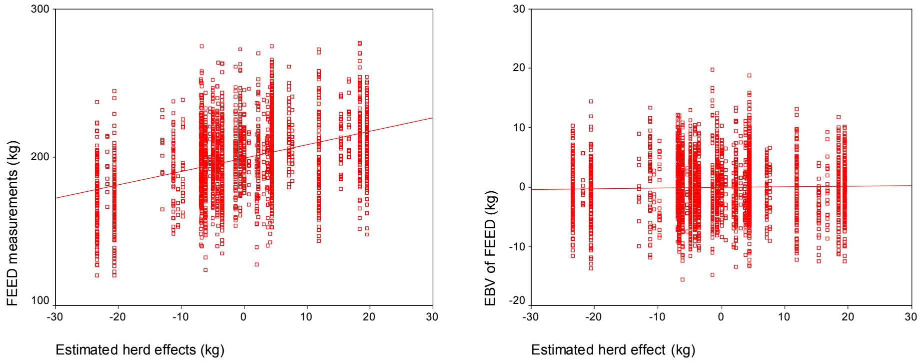

Fig. 5. Scatter plot of the Hungarian Landrace test animals’ FEED measurements and predicted breeding values (EBV) according to the test animals’ herds of origin

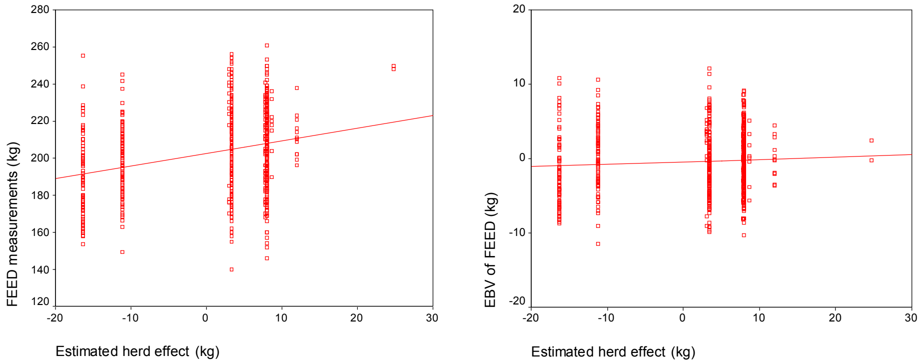

Fig. 6. Scatter plot of the Ka-Hyb hybrid (A-line) test animals’ DOT measurements and predicted breeding values (EBV) according to the test animals’ herds of origin

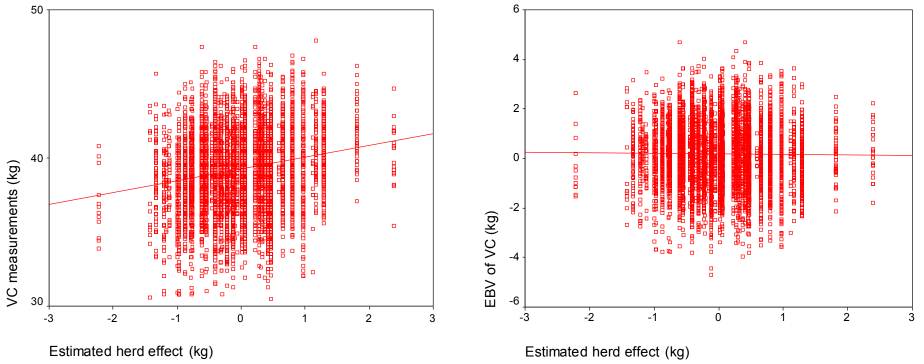

Fig. 7. Scatter plot of the Hungarian Large White test animals' VC measurements and predicted breeding values (EBV) according to the test animals' herds of origin

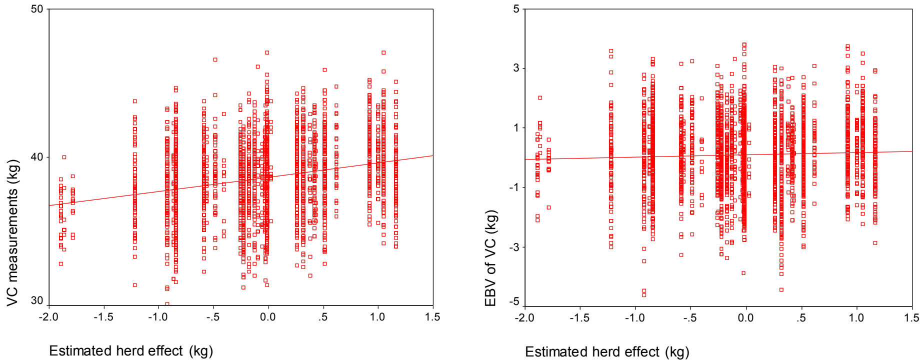

Fig. 8. Scatter plot of the Hungarian Landrace test animals' VC measurements and predicted breeding values (EBV) according to the test animals' herds of origin

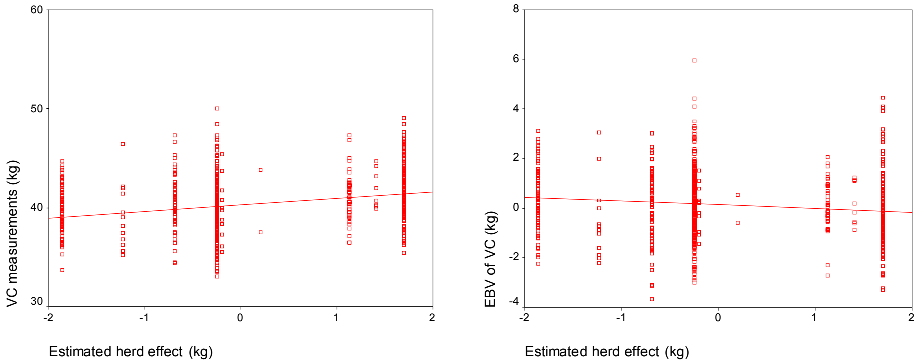

Fig. 9. Scatter plot of the Ka-Hyb hybrid (A-line) test animals' VC measurements and predicted breeding values (EBV) according to the test animals' herds of origin

Concerning the other traits of the test, the smaller amount of feed is consumed (FEED) until reaching the target weight of 105 kg the more favourable the measurement is. That means that, in the same way as it was in case of the DOT, decreasing figures are more advantageous. On the contrary, in case of VC increasing numbers are more favourable. Nevertheless receiving similar results with FEED and VC (Figs 4–9, Tables 3–4) to those of DOT it was obvious that those animals reared in good environments generally consumed less feed and provided greater amount of valuable cuts than their poorly reared counterparts. Yet the predicted breeding values and the contemporary group effects generally did not show correlation coefficients significantly different from zero, hence the genetic merit was probably independent of the environmental factors for these traits as well. It has to be mentioned though that in case of MLW the BLUP method slightly overcompensated the herd effects of the VC, which result might be caused by the relatively small data set.

Looking at the results by other authors, Ugarte et al. (1992) approached the same problem in a different way by not analysing real data sets but investigating artificial databases and the consequences of their observable random or nonrandom distribution. Using computer simulation authors intentionally created artificial dairy cattle data sets either with random or nonrandom distribution of animals concerning their breeding values across the flocks of origin. Nonrandom association among animals and contemporary groups resulted in a positive correlation between them. At the same time, random assignment of animals to contemporary groups resulted in nearly zero correlation between contemporary group effects and predicted breeding values justifying our basic assumptions. Based on theoretical considerations it was demonstrated by Visscher and Goddard (1993) that zero bias could be found supposing the random distribution of animals concerning their breeding values across the herds regardless of the contemporary groups' status (either fixed or random). These results provide essential guidance for the present situation as zero correlation coefficients suggest random distribution (Ugarte et al., 1992) that makes an unbiased breeding value estimation possible (Visscher and Goddard, 1993) on which the process of selection can be based.

Nevertheless, in case of a nonrandom distribution, bias can be removed treating herd effect as fixed. This conclusion can be shown algebraically (Van Vleck, 1987; Hofer and Frey, 1995). However, Visscher and Goddard (1993) noted that if certain animals have progeny only in better environments than others, their evaluation will be biased upwards if contemporary groups are treated as random effects (i.e. evaluation is unbiased treating the contemporary groups as fixed effects). If, however, genetically superior animals are used in better environments, both models, fixed or random contemporary group effects, yield biased evaluations. The results of the present study suggest that this is not the case and therefore the predicted breeding values can be regarded as unbiased.

Conclusions

The present station test index is based on purely phenotypic values. Those animals which have been reared in good environmental conditions prior to the test are therefore clearly favoured. The BLUP procedure may eliminate this source of bias, as a probable random distribution of animals across the flocks according to their genetic merit was found. In the models applied by the present study herd effects were considered as fixed effects; hence, even if nonrandom distribution of animals had occurred across the contemporary groups (which did not seem to be the case), the predicted breeding values would still have been unbiased unless genetically superior animals originated in herds providing favourable environmental conditions.

Acknowledgement

Financial support from the Hungarian Scientific Research Fund (OTKA), project no. TO35151, is acknowledged.

References

Groeneveld, E. (1990): PEST Users' Manual. Institute of Animal Husbandry and Animal Behaviour Federal Research Centre, Neustadt. pp. 1–80.

Groeneveld, E. (1998): VCE4 Users' Guide. Institute of Animal Husbandry and Animal Behaviour Federal Research Centre, Neustadt. pp. 1–61.

Groeneveld, E., Csató, L., Farkas, J. and Radnóczi, L. (1996): Joint genetic evaluation of field and station test in the Hungarian Large White and Landrace populations. Arch. Tierz. **39,** 513–531.

Hazel, L. N. (1943): The genetic basis of constructing selection indexes. Genetics **28,** 476–490.

Hofer, A. (1995): On problems related to data structures and models in genetic evaluation. 2nd Workshop on Advanced Biometrical Methods in Animal Breeding, Salzburg, June 12–20.

Hofer, A. and Frey, M. (1995): Prediction of breeding values for litter size in pigs using an animal model with a fixed or a random herd × year effect. 2nd Workshop on Advanced Biometrical Methods in Animal Breeding, Salzburg, June 12–20.

Mrode, R. A. (1996): Linear models for the Prediction of Animal Breeding Values. CAB International, Wallingford, Oxon, UK.

OMMI (National Institute for Agricultural Quality Control) (2000): Sertés Teljesítményvizsgálati Kódex (Pig Performance Testing Code). 3rd edition, Budapest.

SPSS Inc. (1999): SPSS for Windows. Version 10.

Ugarte, E., Alenda, R. and Carabano, M. J. (1992): Fixed or random contemporary groups in genetic evaluations. J. Dairy Sci. **75,** 269–278.

Van Vleck, L. D. (1987): Contemporary groups for genetic evaluations. J. Dairy Sci. **70,** 2456–2464.

Visscher, P. M. and Goddard, M. E. (1993): Fixed and random contemporary groups. J. Dairy Sci. **76,** 1444–1454.