Effect of De-Regression on Lactation Number on MACE Evaluations: A Simulation Study

Fabiola Canavesi and Filippo Miglior
ANAFI, Cremona, Italy

Introduction

National evaluation procedures applied in dairy cattle populations differ across countries in several ways. Different models are used in various countries (sire, animal, lactation or Test-Day model). Adjustments for fixed environmental effects such as age at calving, month of calving, parity are made using different approaches. Heterogeneity of variance across herds is accounted for either by pre-adjusting the data or directly in the model. Also, differences are found regarding number of lactations used per cow in evaluations. Emanuelson (1999) recently reported that ten countries, out of the twenty-two countries that participated in the International Evaluation of February 1999, used the first three lactations only, and five used the first 5 lactations, while four countries used all lactations available for each cow.

Results from national evaluations are combined with a multi-trait sire model to provide international genetic evaluation of bulls. A model proposed by Schaeffer (1994) called MACE (Multiple Across Countries Evaluation) is currently applied at the Interbull Centre for routine international genetic evaluation of dairy bulls, both for production and type traits. For all traits and for practical reasons the actual number of daughters is used to weigh national bull proofs (Banos and Sigurdsson, 1996), although the actual observations used to regress proofs may vary from one to several lactations, or to several test-day records. Number of daughters per bull along with all pedigree information among bulls are thus used in the de-regression step that leads to genetic parameters estimation (within country sire variance and genetic correlations). These parameters are used in the multi-trait evaluation to compare bulls across countries.

Alternative weighting factors were tested in a simulation by Fikse and Banos (1998, 1999) using a relatively simple population structure in which contemporary group size was set to 10 or 20. They concluded that the use of different weighting factors had an impact on sire variance estimates.

The objective of the present study is to investigate the use of different weighting factor in de-regression and MACE evaluation, taking into account number of lactations and using a real data structure.

Materials and Methods

Ten years of lactation records from 1988 onward from two Italian provinces made up the population structure. This provided a realistic data structure both at the herd and pedigree levels The two provinces were Milan and Mantua, Milan representing medium to large farms and Mantua, small to medium farms. Two generations of pedigree were extracted from the National Herd Book data base. Lactations from the two provinces together with pedigree information constituted a population of about 600,000 animals and 1,000,000 records related to approximately 400,000 cows.

Production records were simulated as:

$$Y = \mu + TBV_{anim} + PE_{anim} + error$$

where:

μ	is a general average effect varying according to year of production.
TBV_{anim}	is the true breeding value of the animal
PE_{anim}	is the permanent environmental effect of each cow.

A phenotypic standard deviation of 1,200 kg was simulated with an heritability of 0.25 which sets the genetic standard deviation to 600 kg.

For each ancestor in the base population (animals with unknown parents) the TBV was randomly generated from a normal distribution with mean 0 and a SD of 600 kg. The progeny TBV were then simulated as:

$$TBV_{anim} = ½TBV_{sire} + ½TBV_{dam} + MS$$

where:

MS is the mendelian sampling

Since the two provinces were simulated as one population, the genetic correlation was assumed to be equal to 1.

Data for the two provinces were analyzed separately with an Animal Lactation Model with repeated records. The evaluation for both provinces included all lactations available per cow and also a maximum of three lactations per cow. Heritability was in both cases 0.25.

Results from the two national evaluations were then used as input in an international Mace evaluation using official programs from Interbull in a two country comparison. Three comparisons were made:

1) Milan bull proofs using all lactations (MIALL) were compared to Mantua all lactation bull proofs (MNALL) using actual number of daughters as weighting factor;
2) Milan using three lactation bull proofs (MI3) compared to MNALL using again actual number of daughters;
3) MIALL compared to MNALL using a weighting factor that accounted for data structure in terms of lactations, contemporary group size and mates (Fikse and Banos (1999).

Sire variances estimated were analyzed and estimated BV at the international level were compared to national evaluations and TBV.

Twenty replicates were considered for each simulation.

Results and Discussion

Table 1 shows distribution of sires in MIALL, MNALL and MI3 data. Mantua is fairly small compared to Milan. *Table 2* reports differences of estimated sire variances compared to the true value of 300 kg that was expected, averaged over 20 replicates. Sire variances proved to be sensitive to data structure. The use of actual number of daughters led to over-estimation of sire variances. With the use of a weighting factor that takes into account the structure of the data better than the simple use of number of daughters, estimated values were closer to the true ones but still tended to be over-estimated. This confirms results from Fikse and Banos (1998, 1999) on simulated data well balanced over time.. Alternative weighting may help improve international comparison of bull proofs especially in those populations where data structure is not homogeneous over time. In case of small populations with poor distribution of data, like the province of Mantua in this study, the alternative weighting improved the variance estimations but not as much as in the bigger data set of Milan. Evaluation of bulls based on a maximum of three lactations instead of all lactations available did not confirm previous findings (Canavesi et al., 1999) perhaps due to the small size of the province sampled compared to the complete Italian data base. *Table 3* reports simple and rank correlations of EBV with TBV. While no differences were evident, there was a slight increase in the correlations with TBV of MACE proofs, due to the increased number of information gained from the almost 500 bulls in common between the two data sets.

Conclusions

Use of three lactations in national evaluations had a small impact on sire variance estimates at the international level, contrary to previous results. Further analysis on large and more unbalanced data sets may be required to verify this finding. Use of a weighting factor that accounted for data structure better than the actual number of daughters (officially used in international evaluations) did show an effect on sire variance estimates and reduced the size of the overestimation. Impact on estimated breeding values was not so evident, at least in terms of correlation with true breeding values. Further studies, testing alternative weighting on more complex data structure, are needed to

determine the better factors to be used in the de-regression process.

References

1) Banos, G. & Sigurdsson, A. 1996. Application of contemporary methods for the use of international data in national genetic evaluation. J. Dairy Sci., 79, 1117-1125.
2) Canavesi, F., Samorè, A.B. & Miglior, F. 1999. Three lactations vs. all lactations model. Abstract of 50th Annual Meeting of EAAP, Poster G2.13.
3) Emanuelson, U., Fikse, F. & Banos, G. 1999. Impact on national genetic valuation models on international comparisons. Presented at Computational cattle breeding '99, Finland, March 99.
4) Fikse, W.F. & Banos G. 1999. Weights on daughter information in international genetic evaluations: effect on estimation of sire variances. Presented at Interbull Workshop, Cremona, October 1998.
5) Fikse, W.F. & Banos G. 1998. Weights on daughter information in international genetic evaluations: effects on (co)variance components. ADSA, 1999.
6) Schaeffer, L.R. 1994. Multiple-country comparison of dairy sires. J. Dairy Sci. 77, 2671-2678.

Table 1. Distribution of sires by birth year for the two provinces

Birth year	MIALL	MNALL
< 79	261	210
80	60	39
81	58	44
82	60	58
83	62	39
84	72	43
85	79	42
86	127	73
87	105	50
88	100	61
89	96	53
90	73	44
91	80	25
92	63	18
93	71	32
94	37	15

Table 2. Estimated sire variances (kg)

Weight	MIALL	MNALL	MI3
N daughters, all lactations	30.3 ± 7.7	39.6 ± 8.1	29.9 ± 7.8
DE yield, mate	14 ± 7.3	30.8 ± 7.9	-

Table 3. Estimated correlations with TBV

	Simple	Rank
MIALL	0.89	0.88
MI3	0.89	0.88
MNALL	0.92	0.91
ITBALL	0.90	0.89
ITB3	0.90	0.89
ITBW	0.90	0.89