Multiple-trait genetic evaluation for milk, fat and protein yields and persistency[1]

N. Gengler

FNRS, UER de Zootechnie, Faculté Universitaire des Sciences Agronomiques, B-5030 Gembloux, Belgium.

INTRODUCTION

Single-trait animal-models (ST-AM) [e.g., DUCROCQ, 1990; VANRADEN and WIGGANS, 1991; LEROY *et al.*, 1993] are currently being implemented all over the world for genetic evaluation of milk, fat and protein lactation yields. But such evaluations are only sub-optimal as there are at least three reasons for the use of multiple-trait animal models (MT-AM). First, such methods improve the precision of the evaluations as error variances (PEV) of predictions are reduced [SCHAEFFER, 1984; THOMPSON and MEYER, 1986], therefore gain from selection would be enhanced. Secondly, an MT-AM would reduce selection bias as selection for yield is done in dairy cattle on milk, fat and protein together but genetic evaluations are still single-trait [POLLAK *et al.*, 1984]. MT-AM can be adapted for missing values, therefore the fact that one or two traits are missing can be taken into account. Even if the advantages of multiple-trait models for milk, fat and protein lactation yields are only limited, such methods enable the introduction of other traits in the models, as e.g., somatic cell scores, productive life and persistency of lactation yields.

The major inconveniences of MT-AM are the programming and solving difficulties due to more complicated matrix structures. But recent advances around canonical transformation make new developments possible. SMITH and LIN proposed in 1990 multiple-diagonalization of (co)variance matrices. DUCROCQ and BESBES [1993] found an easy approach to missing values. Approximate reliability estimation for such methods were described by GENGLER and MISZTAL [1995b]. DUCROCQ and CHAPUIS [1995] and GENGLER and MISZTAL [1995b] proposed two different approaches to solve mixed model equations with different models per trait.

MATERIAL AND METHODS

Data were obtained from A. Toussaint, and A. De Bast, ELINFO (Élevage Informatique), Ciney, Belgium. Two data sets were prepared for the same 32185 cows, with a first record known over 250 days in milk. Records were from the Holstein-Friesian and Red and White Breeds and were performed during the first, second or third lactation. The first data set (Data1) included 44057 lactations recorded for milk, fat and protein yields with over 250 days in milk. The second file (Data2) consisted of all first, second and third lactation records (51555) performed by the 32185 cows with a minimum of 100 days in milk. If records were below 250 days, persistency traits were considered missing and records for yields were corrected for lactation length as those records were considered in progress. Persistency was measured as a function of the variation of partial lactation yields. The parts considered were the first, second and third 100 (105) days [GENGLER *et al.*, 1995]. The values obtained, which

---

[1] **INTERBULL Open Meeting, September 7-8, 1995, Prague, Czech Republic**

were expressed intra-lactation in standard-units, were declared apparent persistency as opposed to real persistency that was corrected for the phenotypic influence of total yields. The real persistency of milk, fat and protein yields measures RSP_M, RSP_F and RSP_P were computed.

Herd-year groups had a minimum of 5 animals per group. Twelve age-parity classes were defined based on parity and inside parity on age of calving. Six seasons were defined, every season grouping two month (e.g., January-February). Eight calving interval classes were defined (≤ 10 month, 11 month,..., 16 month, ≥ 17 month). Records that had no known back calving date were assigned to the last class (≥ 17 month).

Pedigree information for a total of 48554 animals, cows and their ancestors, were obtained from A. Toussaint, and A. De Bast, ELINFO (Élevage Informatique), Ciney, Belgium. Genetic groups were created [WESTELL *et al.*, 1988], based on sex, breed and birth year of animal with at least one unknown parent and on sex of missing parent. Two breed types were in the data set Holstein-Friesians and Red and Whites. Animals recorded as Red and White but with 100% Holstein inheritance were considered Holsteins. The preparation left us with 56 genetic groups.

The models used were close to the current ST-AM used in Belgium [FARNIR *et al.*, 1993] and had the general form:

$$\mathbf{y} = \mathbf{Hh} + \mathbf{Xs} + \mathbf{Tc} + \mathbf{Zp} + \mathbf{Z^{*}u} + \mathbf{e}$$

where: $\mathbf{y}$ is a vector of observations; $\mathbf{H}$, $\mathbf{X}$, $\mathbf{T}$, $\mathbf{Z}$ and $\mathbf{Z^{*}}$ are known incidence matrices; $\mathbf{h}$ is a vector of unknown fixed herd-year or management group effects; $\mathbf{s}$ is a vector of unknown fixed age-parity-season effects; $\mathbf{c}$ is a vector of unknown fixed calving interval group effects; $\mathbf{p}$ is a vector of unknown random permanent environment effects; $\mathbf{u} = \mathbf{a} + \mathbf{Qg}$ with $\mathbf{a}$ being the vector of additive genetic effects and $\mathbf{g}$ the vector of genetic group effect, $\mathbf{Q}$ a known incidence matrix.

Three different models were tested:

Model I: ST-AM for the traits milk, fat and protein yield and real persistencies (file Data1);
Model II: MT-AM for the same traits, no missing values (file Data1);
Model III: MT-AM identical to Model II, but missing values (file Data2).

Single-trait evaluations were performed using an iteration on data approach with Jacobi and second-order Jacobi iterations. The program used was the JAA program provided by I. Misztal [MISZTAL and GIANOLA, 1987]. Multiple-trait evaluations used the following steps: a) multiple-diagonalization of covariance matrices using the FG algorithm [FLURY and CONSTANTINE, 1985]; b) transforming of data to canonical scale; c) resolution of equations and PEV estimation as for ST evaluations; d) back-transformation of solutions and PEV [MISZTAL *et al.*, 1993] to original scale. The program used was called MTJAA and its design was based on the JAA program that was used as sub-routine for step c). Therefore the comparison between ST and MT could be done independently from programming aspects. The PEV of animal solutions were estimated indirectly by the method described by MISZTAL and WIGGANS [1988], MISZTAL *et al.* [1991] and MISZTAL *et al.* [1993]. Model III needed additional programming for the implementation of the solving method described by DUCROCQ and BESBES [1993] and the PEV estimation method described by GENGLER and MISZTAL [1995a]. Computations were done on a DECstation 5000-240, no other programs were executed at the same time. The (co)variance components used, were estimated previously using Model II, Data1 and EM-REML. The objective of this study was two fold: to show the advantages and feasability of a multiple-trait approach and this with the example of milk, fat and protein yields and persistency of yields.

RESULTS AND DISCUSSION

Permanent environment correlations, repeatabilities and residual correlations are given in Table 1, genetic correlations, heritabilities and phenotypic correlations are given in Table 2. The values shown are based on estimated (co)variance components that were used for the computations.

Table 1. Permanent environment correlations above, repeatabilities on and residual correlations below the diagonal among milk, fat and protein yields and real milk, fat and protein persistencies (RSP_M, RSP_F and RSP_P).

	Trait					
Trait	Milk	Fat	Protein	RSP_M	RSP_F	RSP_P
Milk yield	0.48	0.73	0.94	0.28	0.16	0.14
Fat yield	0.89	0.46	0.83	0.20	0.19	0.07
Protein yield	0.95	0.90	0.47	0.23	0.13	0.11
RSP_M	-0.10	-0.07	-0.07	0.26	0.62	0.74
RSP_F	-0.04	-0.06	-0.01	0.47	0.15	0.77
RSP_P	-0.02	-0.01	-0.01	0.48	0.71	0.10

Table 2. Genetic correlations above, heritabilities on and phenotypic correlations below the diagonal among milk, fat and protein yields and real milk, fat and protein persistencies (RSP_M, RSP_F and RSP_P).

	Trait					
Trait	Milk	Fat	Protein	RSP_M	RSP_F	RSP_P
Milk yield	0.20	0.68	0.89	0.08	0.01	0.01
Fat yield	0.80	0.19	0.76	0.03	0.09	-0.04
Protein yield	0.93	0.85	0.22	0.01	-0.09	-0.04
RSP_M	0.00	0.00	0.00	0.14	0.81	0.90
RSP_F	0.00	0.00	0.00	0.51	0.06	0.86
RSP_P	0.00	0.00	0.00	0.52	0.72	0.04

Execution times are difficult to estimate because they depend on many factors such as computing platform, programming language, compiler, etc. But relative execution times were as expected with Model II taking only as much time as the six Model I evaluations together. Results for Model III were different as programming strategies were not completely similar and the degree of optimization of this program was inferior compared to the others. But even without additional polishing a genetic evaluation equivalent to those currently done in Belgium would have taken 40 days for all six traits together using Model III and even less if Model II had been used.

Passing from Model I to Model III can be considered an improvement in the precision of animal ranking obtained through the use of a more and more complicated model. If now rerankings are important, the use of a more complicated model can be justified. Rank coefficients were computed between the breeding values for persistency traits obtained with the three models. Results are given for the 32185 cows with records (Table 3) and for the 1059 sires of these cows (Table 4).

Table 3. Comparison of breeding values for persistency of milk, fat and protein yields (RSP_M, RSP_F and RSP_P) for 32185 cows with records obtained from Model I (single-trait) and Models II and III (multiple-trait).

	Trait		
Rank correlations	RSP_M	RSP_F	RSP_P
Model I - Model II	0.984	0.765	0.729
Model I - Model III	0.955	0.760	0.710
Model II - Model III	0.972	0.966	0.966

Table 4. Comparison of breeding values for persistency of milk, fat and protein yields (RSP_M, RSP_F and RSP_P) for 1059 sires of cows with records obtained from Model I (single-trait) and Models II and III (multiple-trait).

	Trait		
Rank correlations	RSP_M	RSP_F	RSP_P
Model I - Model II	0.989	0.791	0.733
Model I - Model III	0.960	0.807	0.744
Model II - Model III	0.969	0.946	0.960

The use of Model II instead of Model I produced an important reranking of animals, especially for fat and protein persistencies. The use of Model III instead of Model II was at the origin of a smaller reranking that was greater for sires. These results can be explained as follows. The additional information provided in Model II comes essentially from other traits. Therefore traits with lower heritabilities are more affected. The additional information provided by records in progress added in Model III comes through correlations with yield traits. Despite low genetic correlations reranking was still important, an indication that consideration of records in progress affects results for persistency traits.

Table 5. Mean prediction error variances (PEV) for single-trait (ST Model I) and multiple-trait (MT Model II) models and relative gain in prediction error variance for genetic evaluation of persistency for milk, fat and protein yields (RSP_M, RSP_F and RSP_P).

	Trait		
	RSP_M	RSP_F	RSP_P
Cows with records			
Mean ST-PEV	0.0680	0.0411	0.0313
Mean MT-PEV	0.0674	0.0383	0.0287
Relative gain in PEV	0.82%	6.97%	8.39%
Sires of cows			
Mean ST-PEV	0.0593	0.0369	0.0284
Mean MT-PEV	0.0591	0.0346	0.0257
Relative gain in PEV	0.43%	6.22%	9.55%

Table 6. Mean accuracy (r_{TI}) for single-trait (ST Model I) and multiple-trait (MT Model II) models and relative improvement of genetic gain due to the use of multiple-trait for genetic evaluation of persistency for milk, fat and protein yields (RSP_M, RSP_F and RSP_P).

	Trait		
	RSP_M	RSP_F	RSP_P
Cows with records			
Mean accuracy-ST	0.502	0.397	0.359
Mean accuracy-MT	0.508	0.473	0.456
Relative genetic gain	1.30%	19.03%	26.97%
Sires of cows			
Mean accuracy-ST	0.564	0.470	0.433
Mean accuracy-MT	0.567	0.520	0.514
Relative genetic gain	0.48%	10.79%	18.76%

The relative gain due to the use of MT instead of ST can be estimated through different approaches. A first method is based on the relative reduction of mean PEV [SCHAEFFER, 1984] due to the use of MT. The reductions are of 0.82% for milk, 6.97% for fat and 8.39% for protein persistencies for the 32185 cows with records and of 0.43%, 6.22% and 9.55% for 1059 sires (Table 5). A second possibility is to assess the improvement due to MT through the increased genetic gain (Table 6). As genetic gains are proportional to accuracy expressed as the correlation between estimates and real breeding values, the ratio of mean accuracy for MT to mean accuracy for ST gives the approximate relative genetic gain. The relative genetic gain for cows was 1.30% for milk, 19.03% for fat and 26.97% for protein persistencies, and 0.48%, 10.79% and 18.76% for sires. This improvement is spectacular, but might be expected as heritabilities for fat and protein persistencies are low, correlations with milk persistencies are always high to very high.

CONCLUSION

Passing from ST-AM to MT-AM for yield traits is possible using canonical transformation. This can be done without great expense. The advantages for yield traits are only limited, but the real advantage of MT is that it permits the easy introduction of other traits in such models. Joint evaluation would result in reduced selection bias and increased precision. Traits that could be analyzed in such models are somatic cell counts (scores), productive life but also the associated yield traits: persistency of yields. This study showed that the MT approach has certain advantages: accommodate missing values due to records in progress and improvement of accuracy of up to 26%. As a final conclusion, one can say that MT-AM is an alternative for current models, as for example in Belgium, but its real interest is the introduction of additional traits.

ACKNOWLEDGEMENTS

The author wishes to thank I. Misztal, University of Illinois, Urbana, USA, for providing his JAA computer program and A. Toussaint and A. De Bast, ELINFO, Ciney, Belgium, for providing the data.. N. Gengler is Aspirant of the Fonds National Belge de la Recherche Scientifique. Computer programs JAA and MTJAA are available electronically from anonymous FTP at 128.174.78.6..

REFERENCES

DUCROCQ V. [1990]. Les techniques d'évaluation génétique des bovins laitiers. INRA *Prod. Anim.* **3**, 3 - 16.

DUCROCQ V., BESBES B. [1993]. Solution of multiple trait animal models with missing data on some traits. *J. Anim. Breed. Genet.* **110**, 81 - 92.

DUCROCQ V., CHAPUIS H. [1995]. Solution of multiple trait mixed model equations: an approach a canonical transformation when different models describe different traits. Annual Meeting of the ADSA, Ithaca, USA, June 25 - 28, 1995.

FARNIR F., VOLCKAERT D., SPEYBROEK E., LEROY P. [1994]. Milk and dual purpose basis for the publication of breeding values estimated with the animal model in Belgian dairy breeds. Open Session of the Interbull Annual Meeting, Ottawa, Canada, August 5 - 6, 1994.

FLURY B. N., CONSTANTINE G. [1985]. Algorithm AS211: the F–G diagonalization algorithm. *Appl. Stat.* **34**, 177 - 185.

GENGLER N., KEOWN J.F., VAN VLECK L.D. [1995]. Various persistency measures and relationships with total, partial and peak yields. *Br. J. Genetics* **18**, 237 - 243.

GENGLER N., MISZTAL I. [1995a]. Approximation of reliability for multiple–trait animal models with missing data. *J. Dairy Sci.* (accepted).

GENGLER N., MISZTAL I. [1995b]. A simple algorithm to solve multitrait mixed linear models with different fixed effects per group of traits using canonical transformation. 46th Annual Meeting of the EAAP, Prague, Czech Republic, September 4 - 7, 1995.

LEROY P., FARNIR F., VOLCKAERT D. [1993]. Estimation of breeding values of Belgian Black and White, Red and White, White Red and Red of Flanders breeds with the animal model. Open Session of the Interbull Annual Meeting, Aarhus, Denmark, August 19 - 20, 1993.

MISZTAL I., GIANOLA D. [1987]. Indirect solution of mixed model equations. *J. Dairy Sci.* **70**, 716 - 723.

MISZTAL I., LAWLOR T.J., SHORT T.H. [1993]. Implementation of single– and multiple–trait animal models for genetic evaluation of Holstein type traits. *J. Dairy. Sci.* **76**, 1421 - 1432.

MISZTAL I., WIGGANS G.R. [1988]. Approximation of prediction error variance in large–scale animal models. *J. Dairy Sci.* **71**(Suppl. 2), 27 - 32.

MISZTAL I., WIGGANS G.R., LAWLOR T.J., SHORT T.H. [1991]. Continuous genetic evaluation of dairy cattle. *J. Dairy Sci.* **74**, 2001 - 2011.

POLLAK E.J., VAN DER WERF J., QUAAS R.L. [1984]. Selection bias and multiple trait evaluation. *J. Dairy Sci.* **67**, 1590 - 1595.

SCHAEFFER L.R. [1984]. Sire and cow evaluation under multiple trait models. *J. Dairy Sci.* **67**, 1567-1580.

THOMPSON R., MEYER K. [1986]. A review of theoretical aspects in the estimation of breeding values for multi-trait selection. *Livest. Prod. Sci.* **15**, 299 - 313.

VANRADEN P.M., WIGGANS G.R. [1991]. Derivation, calculation, and use of national animal model information. *J. Dairy Sci.* **74**, 2737 - 2746.

WESTELL R.A., QUAAS R.L., VAN VLECK L.D. [1988]. Genetic groups in an animal model. *J. Dairy Sci.* **71**, 1310 - 1318.