

Genetic Correlation Estimates of a Multiple Lactation Multiple Country Model for Milk Production Traits Based on Performance Records

Z. Liu, F. Reinhardt and R. Reents
VIT, Heideweg 1, D-27280 Verden/Aller, Germany

Introduction

Current international bull comparisons are based on estimated breeding values (EBV) from national genetic evaluations. It has been shown that the present method of de-regressing national EBV in the procedure of international genetic evaluation caused bias in case of national evaluations using multiple trait models (Madsen *et al.*, 2001). Though a theoretical extension of the de-regression method to the multiple trait case was outlined (Schaefer, 2001), a practical implementation of the multi-trait de-regression method would be very difficult at the international level, because, among other technical problems, reconstruction of bulls' diagonal blocks in original national mixed model equation systems requires too much information than countries can easily provide, particularly for countries using multiple lactation random regression test day models. In contrast to de-regressed proofs, daughter yield deviations (DYD) can be more readily extended to the models of multiple correlated lactations within countries. Since DYD are unregressed phenotypic measure of bull's genetic merit, de-regression process at the international level will no longer be needed. A key issue in using DYD of bulls in international evaluations is the calculation of DYD and weights for DYD, which is not straightforward for national genetic evaluations using a multiple lactation random regression test day model. An implementation of multiple lactation multiple across country evaluation for milk production traits using lactation specific DYD requires accurate estimates of genetic correlations between lactations across countries, assuming that there are no residual correlations between lactation specific DYD within country (Sullivan & Wilton, 2001). The objective of this study was to estimate genetic correlations between lactations across countries for milk production traits based on pre-corrected 305-day lactation records.

Material and Methods

A multiple lactation multiple country model was applied to lactation yield deviations (LYD) of milk production traits of Holstein cows from France, Germany, Italy and The Netherlands involved in the PROTEJE project (Canavesi *et al.*, 2001). LYD from first three lactations were adjusted for all fixed effects and heterogeneous herd variances on national level in a similar way as described in Ducrocq *et al.* (2001), and accuracy associated with the LYD was also provided for individual lactations. For the countries using a repeatability animal model in national genetic evaluations, LYD were back transformed to original scale from either Mature Equivalents (ME) or Two-Years-Old Equivalents (TE) scale. A sire model was chosen for estimating genetic parameters of first three lactations of the four countries:

$$y_{ijkl} = \mu_{ij} + s_{ijk} + e_{ijkl}$$

where y_{ijkl} is LYD of lactation j of cow l of sire k in country i , μ_{ij} is general mean of lactation j in country i , s_{ijk} is effect of sire k of lactation j in country i , and e_{ijkl} is residual effect. Accuracy of y_{ijkl} is denoted as R_{ijkl}^2 .

Since the identity of the participating countries is irrelevant for the present study which emphasizes a theoretical development for international genetic evaluation, the four participating countries were randomly labelled as A, B, C and D throughout the paper. Original LYD records were edited on lactation level according to the criteria: parity number between 1 and 3; age at calving between 21 and 70 months; days in milk of last test between 240 and 330; and calving years between 1990 and 1997 for first lactation, between 1991

and 1997 for second lactation and between 1992 and 1997 for third lactation, respectively. Lactation records from other than the monthly milk testing programme were deleted. Cows were required to have first lactation records and known sires. No cows were allowed to have data in more than one country to avoid estimation of residual correlations between countries. Bulls tested in only one of the four participating countries must have at least 30 daughters in order to be selected for further analyses. Daughters in foreign countries of own bulls and all daughters of bulls from countries other than the participating ones were kept to achieve sufficient direct ties across countries. International bull pedigree and cross reference files from Interbull November 2001 evaluation were used to trace back all ancestors of the bulls through sires and maternal grandsires. No phantom groups were assigned to unknown parents. Table 1 summarises the structure of the

data set after all edits and phenotypic standard deviations of LYD of first three lactations for all three yield traits. Approximately 8.9 million lactation records from c.a. 4.6 million cows remained for the parameter estimation. Since a back transformation from ME to original scale was not possible for country C as a result of lack of necessary information, LYD of the three lactations are more similar in standard deviations than data from country A or B. For national genetic evaluations, country D used TE factors that differ by month of age, thus a complete restoration of LYD to their original scales does not seem to be successful, which is reflected in the less differentiated standard deviations of LYD between the lactations. It can be seen in Table 2 that the data sets from the four countries have sufficient direct ties through common sires of cows as well as sufficient indirect ties through common ancestors of the bulls.

Table 1. Descriptive statistics of the final data set for parameter estimation.

Country	No. of lactations	No. of cows	Standard deviations of milk / fat / protein yield (kg)		
			Lactation 1	Lactation 2	Lactation 3
A	2,872,941	1,443,663	970 / 38 / 28	1195 / 47 / 35	1180 / 48 / 35
B	2,613,598	1,455,714	1047 / 41 / 30	1210 / 49 / 35	1267 / 54 / 37
C	1,458,831	760,025	1277 / 49 / 37	1323 / 53 / 39	1310 / 52 / 38
D	1,954,527	924,894	1061 / 41 / 31	1111 / 44 / 33	1094 / 45 / 33
Total	8,899,897	4,584,296			

Table 2. Number of common bulls and the size of the final pedigree file.

Sires of cows in	No. of sires of cows in country				No. of sires of cows	No. of animals in sire pedigree
	A	B	C	D		
One country	3,809	4,783	2,333	1702	12,627	
Two countries	1,725	574	734	1457	2,245	
Three countries	506	286	429	489	570	
Four countries	409	409	409	409	409	
Total	6,449	6,052	3,905	4057	15,851	16,883

Residual maximum likelihood estimates of sire as well as residual (co)variances were obtained by applying VCE package (Neumaier & Groeneveld, 1998) separately for one of the three yield traits. Estimating 102 (co)variance components and solving a mixed model equations with a size of 202,600 equations were computationally very challenging. Therefore, four sub-analyses with data from three countries each were performed. A simple average of the estimates from the sub-analyses was then calculated and regarded as final estimate of the parameters. The parameter estimates from the sire model were converted to

the animal model base. As the VCE package cannot analyse records with variable weights or weights less than one, it was assumed throughout the parameter estimation that $R_{ijkl}^2 = 1$.

Results and Discussion

Fortran 90, SAS and Maple 6 programs were developed for the research project. The parameter estimation was conducted on an Unix server HP9000/L2000. Estimated heritabilities, genetic

and phenotypic correlations were given in Table 3, 4, and 5 for milk, fat and protein yield, respectively.

Within country parameter estimates: In diagonal blocks corresponding to individual countries are heritability and genetic correlation estimates of first three lactations for the respective countries. The estimates should be in close agreement with parameters used in national genetic evaluations. Compared to the heritability estimates by Weigel *et al.* (2001) for first lactation production traits,

the heritability estimates of 305-day LYD are much higher, which may be explained by the fact, among others, that the pre-corrected lactation records were assumed to have an unity reliability in the parameter estimation. This may have caused an underestimation of residual variances and thus an overestimation of heritability. For international evaluations enabling multiple lactations per bull per country national parameters of first three lactations from participating countries should be used.

Table 3. Estimated genetic correlations (above diagonal), heritabilities (on diagonal) and phenotypic correlations (below diagonal) for 305-day lactation milk yield.

Country / Lactation		A			B			C			D		
		1	2	3	1	2	3	1	2	3	1	2	3
A	1	.57	.90	.91	.88	.78	.77	.91	.85	.82	.92	.87	.85
	2	.60	.45	.96	.74	.81	.82	.82	.87	.84	.79	.87	.85
	3	.56	.57	.45	.84	.87	.89	.88	.92	.92	.84	.91	.92
B	1	.12	.09	.10	.49	.90	.88	.91	.85	.83	.89	.83	.83
	2	.10	.09	.10	.51	.42	.98	.86	.90	.89	.84	.90	.90
	3	.09	.09	.10	.45	.53	.39	.85	.91	.91	.83	.91	.92
C	1	.13	.10	.11	.12	.10	.10	.55	.96	.94	.91	.89	.88
	2	.11	.10	.11	.11	.10	.10	.61	.48	.99	.85	.91	.91
	3	.10	.09	.10	.10	.10	.09	.56	.61	.43	.83	.90	.92
D	1	.15	.11	.12	.14	.12	.11	.15	.13	.12	.59	.95	.93
	2	.13	.12	.12	.12	.12	.11	.13	.13	.12	.66	.52	.99
	3	.13	.11	.12	.11	.12	.11	.13	.12	.12	.61	.66	.50

In general, first lactation has the highest heritability among the three lactations, which is observed in all countries for all traits. Among the three production traits, heritability estimates are the highest for milk yield, followed by protein and fat yield, and further the differences in heritability estimates between protein and fat yields are much smaller than between milk yield and either of them. Compared to countries A and B, heritability estimates of countries C and D are even more higher. The reason for high heritability estimates for country C may be that the analysed LYD were

on ME basis, instead of original scale, because of a lack of information for a back transformation of the data. For the high heritabilities for country D one may argue that the seemingly unsuccessful back transformation of LYD from TE, which differ by month of age, is responsible for the high heritability estimates as it is the case for country C. It is noteworthy that residual correlation between any two lactations is slightly negative for country D (not shown here), but all positive for the other three countries.

Table 4. Estimated genetic correlations (above diagonal), heritabilities (on diagonal) and phenotypic correlations (below diagonal) for 305-day lactation fat yield.

Country / Lactation		A			B			C			D		
		1	2	3	1	2	3	1	2	3	1	2	3
A	1	.44	.90	.89	.90	.81	.80	.91	.85	.84	.94	.87	.85
	2	.59	.37	.97	.78	.88	.88	.87	.90	.90	.84	.93	.92
	3	.53	.57	.37	.82	.90	.92	.87	.91	.93	.84	.91	.93
B	1	.10	.08	.08	.42	.89	.86	.91	.84	.82	.91	.82	.81
	2	.08	.08	.08	.48	.37	.98	.86	.89	.89	.86	.93	.93
	3	.08	.08	.08	.42	.52	.35	.87	.91	.91	.83	.91	.93
C	1	.11	.10	.10	.11	.10	.10	.56	.97	.95	.88	.85	.85
	2	.10	.10	.10	.10	.10	.10	.62	.49	.99	.82	.88	.88
	3	.10	.10	.10	.10	.09	.09	.58	.63	.46	.80	.88	.89
D	1	.12	.10	.10	.11	.10	.09	.12	.11	.10	.55	.93	.91
	2	.10	.10	.10	.10	.10	.10	.11	.11	.11	.63	.49	.99
	3	.10	.10	.10	.09	.10	.09	.11	.10	.10	.57	.63	.45

Table 5. Estimated genetic correlations (above diagonal), heritabilities (on diagonal) and phenotypic correlations (below diagonal) for 305-day lactation protein yield.

Country / Lactation		A			B			C			D		
		1	2	3	1	2	3	1	2	3	1	2	3
A	1	.44	.91	.88	.88	.79	.75	.91	.84	.81	.91	.85	.82
	2	.60	.39	.96	.78	.85	.85	.87	.90	.88	.84	.91	.89
	3	.52	.59	.37	.83	.88	.91	.88	.93	.93	.83	.91	.93
B	1	.10	.08	.09	.45	.91	.85	.89	.83	.81	.87	.81	.79
	2	.08	.08	.09	.50	.39	.98	.86	.89	.88	.82	.89	.89
	3	.08	.08	.09	.41	.54	.36	.83	.89	.90	.78	.88	.90
C	1	.11	.10	.10	.11	.10	.09	.55	.96	.94	.89	.87	.86
	2	.10	.10	.10	.10	.10	.10	.62	.50	.99	.82	.89	.90
	3	.09	.09	.10	.09	.09	.09	.55	.63	.45	.79	.88	.91
D	1	.12	.10	.10	.11	.10	.09	.13	.11	.11	.57	.94	.91
	2	.11	.11	.10	.10	.10	.10	.12	.12	.11	.64	.52	.99
	3	.10	.10	.10	.10	.10	.10	.12	.11	.11	.57	.65	.48

As expected, the lowest genetic correlation was obtained between first and third lactations, followed by the genetic correlation between first and second lactations. The genetic correlation between second and third lactation is very high. The genetic correlation estimates do not differ markedly between traits. Since LYD of all three lactations are on ME basis for country C, its genetic correlations are the highest among the countries, nevertheless the correlations between first and later lactation are obviously less than one. For the same reason, genetic correlations between lactations are also higher for country D than those for countries A and B. Compared to genetic correlation estimates, estimated phenotypic correlations between lactations exhibit similar pattern across lactations and across the

countries, except that the phenotypic correlations are much lower than genetic ones.

Across country parameter estimates: Genetic correlations between the countries were estimated based on direct ties due to daughters in different countries of bulls as well as indirect ties due to common ancestors of bulls (see Table 2). As the four countries are well connected through both types of ties, the genetic correlation estimates should have a high precision, though there were no standard errors reported for the estimates.

Among the three traits, fat yield has the highest correlations between the countries, except between country C and D. In contrast to fat yield, the genetic correlations of milk and protein yield

depend on country pairs. Milk yield has lower correlations for some country pairs but higher for other pairs than protein yield.

For a given pair of countries genetic correlations between the same lactation are always higher than correlations between two different lactations for first or third lactation, but not always for second lactation. For a given lactation in a specific country, its genetic correlation with a different lactation in the same country is higher than its correlation with the different lactation in any of the other countries. Off-diagonal block between any two countries represents a matrix containing nine genetic correlations between lactations for the country pair. It is interesting to note that such matrices of genetic correlations are not symmetric in most cases. For instance, first lactation protein yield in country B has a genetic correlation of 0.83 with third lactation in country A, whereas the genetic correlation between first lactation in country A and third lactation in country B is only 0.75. Such asymmetric matrices of genetic correlations reveal complex structure of genetic correlations between lactations across countries.

Interbull genetic correlation estimates using a single trait multiple country model are expressed on a combined lactation basis, whereas the correlation estimates reported here are those on individual lactation basis. Since lactations are positively correlated with each other, the Interbull correlation estimate should be higher than the average of the correlation estimates between individual lactations for a given pair of countries, which is confirmed in Tables 3, 4, and 5. Compared to the Interbull correlations between countries, the presented genetic correlation estimates are much more variable, with the lowest value of 0.74 between second lactation milk yield of country A and first lactation of country B and the highest value of 0.94 between first lactation fat yields in country A and D. The larger variation in genetic correlations may improve international genetic evaluations in that the genetic correlation matrix will be easier to be estimated and mixed model equations will be easier to be solved. For the studied countries, the genetic correlation estimates of first lactation production traits are in close agreement with the estimates obtained by Weigel *et al.* (2001).

Phenotypic correlations are generally more similar than genetic correlations between lactations and across countries. The differences in phenotypic correlations between traits are minimal.

Summary

A multiple lactation multiple country model was applied to dairy production trait data from four European countries participating in the PROTEJE project in order to estimate genetic correlations between lactations across countries. Instead of original 305-day lactation records, lactation yield deviations which were corrected for all fixed effects and heterogeneous herd variances on national levels were chosen as traits for the PROTEJE project in order to avoid re-estimation of fixed effects and re-adjustment of heterogeneous herd variance on the international level. Approximately 8.9 million lactations from c.a. 4.6 million Holstein cows sired by a total of 15,851 Holstein bulls were kept for parameter estimation after all necessary edits. Both direct ties by common sires of cows and indirect ties by common ancestors of the sires were sufficient for a high precision of the parameter estimates. In contrast to reasonable genetic correlation estimates between lactations corresponding to each individual countries, heritability estimates of individual lactations were very high, in particular for the countries using complex ME or TE system. Among other reasons, the high heritability estimates were caused by ignoring the non-unity weights of the pre-corrected lactation yield deviations in the parameter estimation. A complex structure of genetic correlations between lactations across countries was shown in the across country parameter estimates. Compared to Interbull genetic correlation estimates resulting from a single trait MACE model, the estimated across country genetic correlations were much more variable, which may improve international genetic evaluations in that the genetic correlation matrix would be less likely to be ill-conditioned. International bull comparison for dairy production traits can be improved by using a multiple lactation multiple country model.

Acknowledgements

The participating institutions of the PROTEJE project: ANAFI of Italy, INRA and SGQA of France, Gembloux University of Belgium, and CR-Delta of The Netherlands, are kindly thanked for the collaboration.

References

- Canavesi, F., Boichard, D., Ducrocq, V., Gengler, N., De Jong, G. & Liu, Z. 2001. Production traits European joint evaluation (PROTEJE). *Interbull Bulletin* 27, 32-34.
- Ducrocq, V., Boichard, D., Barbat, A. & Larroque, H. 2001. Implementation of an approximate multi-trait BLUP evaluation to combine production traits and functional traits into a total merit index. *52nd Annual Meeting of the European Association for Animal Production*, Budapest, Hungary.
- Madsen, P., Sørensen, M.K. & Mark, T. 2001. Validation and comparison of methods to estimate (co)variance components for MACE. *Interbull Bulletin* 27, 73-79.
- Neumaier, A. & Groeneveld, E. 1998. Restricted maximum likelihood estimation of covariances in sparse linear models. *Genet. Sel. Evol.* 30, 3-26.
- Schaeffer, L.R. 2001. Multiple trait international bull comparison. *Livest. Prod. Sci.* 69, 145-153.
- Sullivan, P. & Wilton, J.W. 2001. Multiple-trait MACE with a variable number of traits per country. *Interbull Bulletin* 27, 68-72.
- Weigel, K.A., Rekaya, R., Zwald, N.R. & Fikse, W.F. 2001. International genetic evaluation of dairy sires using a multiple-trait model with individual animal performance records. *J. Dairy Sci.* 84, 2789-2795.