Simplification of the MACE Procedures – Same Deregression for Breeding Values and Correlations

Jette H. Jakobsen[1] and Gerald Jansen[2]

[1]Interbull Centre, Department of Animal Breeding and Genetics, SLU, S-750 07 Uppsala, Sweden

[2]Dekoppel Consulting, Chiaverano(TO), Italy

1. Introduction

Currently, a pedigree with sire-maternal grandsire relationship is used for international genetic evaluations. Deregression and sire variance estimation is done twice (see Figure 1) – once for correlation estimation with a data cut off in 1970 and once for breeding value prediction with a data cut off in 1986 (HOL) and 1981 (other breeds). During the revisit of the evaluation process it was proposed to do the deregression only once for both correlation estimation and breeding value prediction.

The current evaluation process is, however, under change as it is known that international genetic evaluations are sensitive to genetic groups. De Jong (2003) suggested including also pedigree on bull dams in order to move genetic groups further away from animals with data. Based on this suggestion a MACE pilot study using sire-dam (SD) genetic relationship was performed by Van der Linde *et al.* (2005). They estimated genetic correlations for seven Holstein populations for protein yield and found very similar genetic correlations using SD-pedigree as the ones estimated using a sire-MGS pedigree while the computing time increased with a factor 28. They also predicted breeding values and found an increase in predictability of proofs when changing pedigree and an increase in computing time with a factor nine. Based on these findings the Interbull Technical Committee recommended to use the SD-pedigree for breeding value prediction but to leave the correlation estimation as it is with the sire-MGS pedigree structure. The new workflow is illustrated in Figure 1.

After the change in workflow to different pedigree structures for correlation estimation and for breeding value prediction the original idea of one de-regression, only, is no longer as obvious. The aim of this study was therefore to identify the impact on genetic correlations when using same SD de-regressed files for breeding value prediction and correlation estimation.

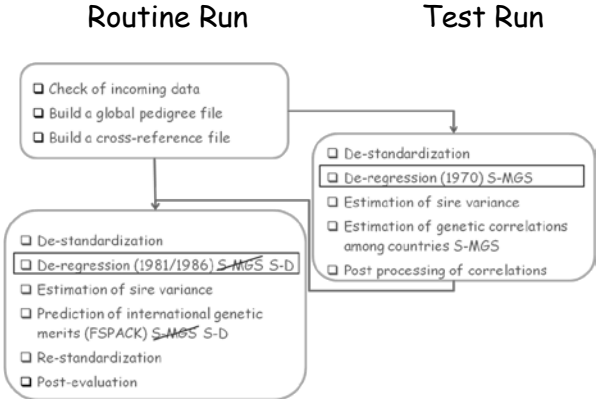

Figure 1. Schematic illustration of the process around routine and test runs.

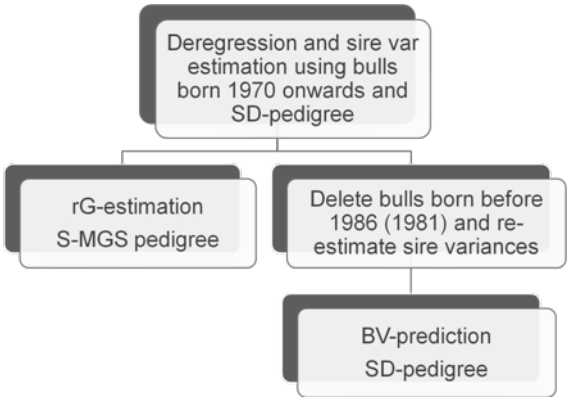

Figure 2. Design for workflow using one de-regression only.

2. Material and Methods

2.1 Pedigree Data

Sire-dam pedigree was extracted from the Interbull database February 5 2010. Pedigree was traced as far back as possible starting from the Sire-MGS pedigree-file used for the January 2010 evaluation and converted to Sire-Dam pedigree format. The pedigree was complemented with pedigree information as provided for the January 2010 routine evaluation but not yet in the database. For the current study, parents of animals born before 1960 were set missing. The Holstein (HOL) pedigree contained 552,998 animals with either sire or dam known and the Red Dairy Cattle (RDC) pedigree 86,872 animals with either sire or dam known.

2.2 National predicted genetic merits

National predicted genetic merits were the same as provided by national evaluation centers and used for the international genetic evaluation in January 2010. The amount of data used for the study (data cut off year 1970) as well as the corresponding heritabilities as provided by the national genetic evaluation centers are shown in Table 1 for protein yield, somatic cell count, direct calving ease, and female fertility traits (interval trait; it) for HOL and for RDC.

2.3 Evaluation steps

The proposed design is illustrated in Figure 2 and described below in Step 1 to Step 4.

Step 1. De-regression and sire variance estimation using the SD-pedigree (described under point 2.1) and PGM (described under point 2.2).

Step 2. Use the de-regressed files and the sire variances computed under *Step 1* and the sire-MGS pedigree as formed for the January routine evaluation and estimate genetic correlations.

Step 3. Delete bulls born before 1986 (HOL) and 1981 (other breeds) from the deregressed files and re-estimate sire variances on the reduced data

Step 4. Predict breeding values using files and parameters from *Step 3* and SD-pedigree

For the current study only *Step 1* and *Step 2* were performed and in addition correlation estimation using the traditional procedure (illustrated under Test Run in Figure 1). In order to check if the size of heritability, the connectedness among countries or sub-setting – no-sub-setting affects the correlations when working with one deregression only the study was performed for high and low heritability traits (protein yield (pr), somatic cell (sc), direct calving ease (dc) as well as female fertility (it)), for breeds with different connectedness (HOL and RDC) and for breed-traits where sub-setting usually is used (HOL-pr; HOL-sc) and where sub-setting is not used (HOL-dc; HOL-it; RDC-pr, RDC-sc; RDC-dc; RDC-it.

2.4 Methods

The HA-USA software (Klei, 1998; Klei & Weigel, 1998) modified by Van der Linde *et al.* (2005) to handle SD-pedigree was used for data preparation. The software was further speed optimized (Jakobsen & Fikse, 2009) including module for de-regression based on Sigurdsson & Banos (1995) and module for sire variance estimation as described by Sullivan (1999) and used in the current study for de-regression and sire variance estimation (*Step 1*). The HA-USA software was used for the correlation estimation using S-MGS-pedigree.

Table 1. Number of records included per country and breed for protein yield, somatic cell, direct calving ease, and female fertility (interval trait).

	Protein Yield		Somatic Cell		Direct Calving Ease		Fertility (it)	
Country	HOL	RDC	HOL	RDC	HOL	RDC	HOL	RDC
CAN	9091	618	9016	617	9677	272	4738	304
DEU	21807	337	21764	337			19204	265
DFS	10575	6112	11682	8025	11544	5536	13585	8612
FRA	19118		14741		7925			
ITA	7813		7967		8043		7427	
NLD	13500		13267		11728		12603	
USA	37708	550	30756	450	29151		38073	514
CHE	915		1098		715			
GBR	9280	726	5659	364	1483		6319	342
NZL	6643	1105	6288	1168			6714	1262
AUS	7350	635	6272	561	763			
BEL	1051		959				888	
IRL	1936		1843				2173	
ESP	2714		2398				1883	
CZE	2762		2445					
SVN	260							
EST	626	293	812	331				
ISR	1162		1034		296			
CHR	1425		1468		733			
FRR	277		209					
HUN	3015		2450		1395			
POL	6099						3248	
ZAF	1570	140	1379	189				
JPN	4728		4634					
LVA	453	599						
SVK	788		830					
NOR		4209		4188		2490		3889
DNR			306					
LTU	351	194						
No. Records	173017	15518	149277	16230	83453	8298	116855	15188

Table 2. Range of heritabilities for protein yield, somatic cell count, direct calving ease, and female fertility (interval trait) for Holstein (HOL) and Red Dairy Cattle (RDC).

Trait	Range of heritability	No of countries (HOL)	No of countries (RDC)
Protein Yield	0.14 – 0.51	27	12
Somatic Cell Count	0.06 – 0.43	24	10
Direct Calving Ease	0.02 – 0.28	12	3
Female Fertility (interval trait)	0.03 – 0.15	12	7

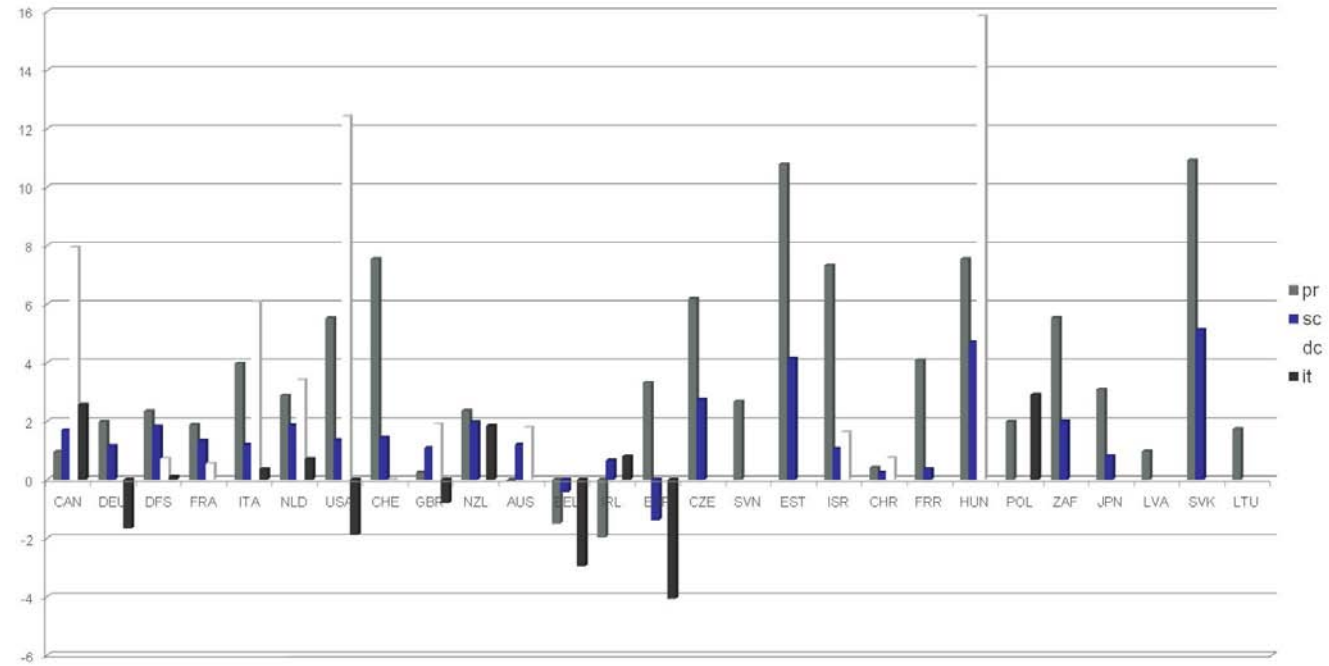

Figure 3. Changes in sire standard deviation in percent when changing from S-MGS pedigree to S-D pedigree for the *Holstein* breed.

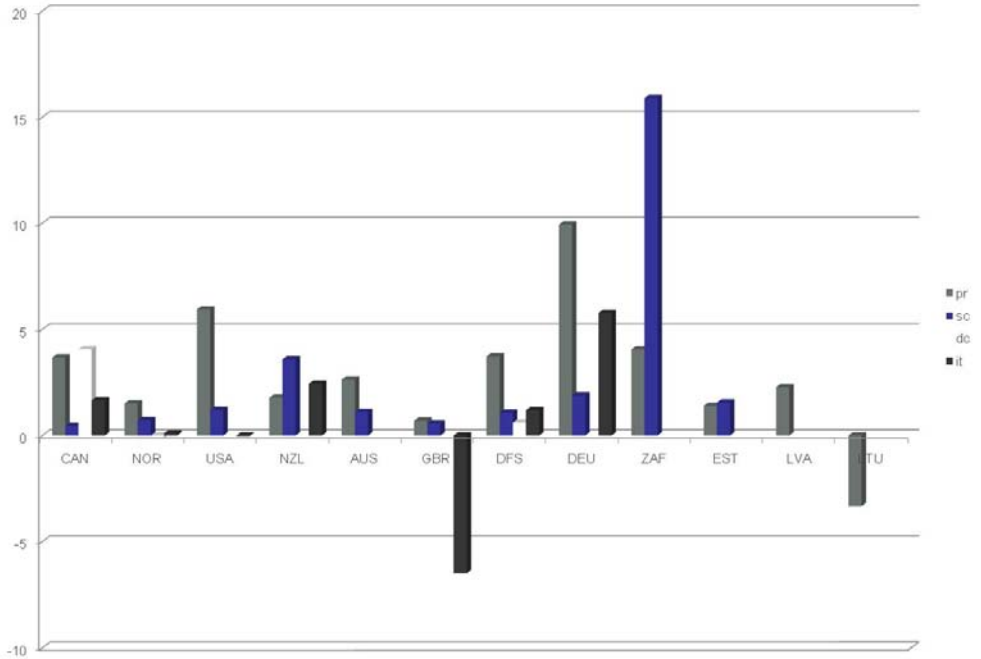

Figure 4. Changes in sire standard deviation in percent when changing from S-MGS pedigree to S-D pedigree for the *Red Dairy Cattle* breed.

Table 3. Average absolute difference, minimum (Min), maximum (Max) and average (Avr) MACE correlations for **Holstein** for protein, cell count, calving ease and female fertility (it) for correlation estimation using S-D deregressed files and S-MGS deregressed files as input for estimation of MACE correlations.

Trait	Avr abs difference	De-regressed SD-files			De-regressed S-MGS files		
		Min	Max	Avr	Min	Max	Avr
Protein	0.024	0.148	0.944	0.722	0.123	0.941	0.709
Cell count	0.009	0.661	0.976	0.884	0.670	0.977	0.887
Calving ease	0.051	0.073	0.977	0.728	0.105	0.963	0.758
Fertility (it)	0.015	0.488	0.972	0.778	0.507	0.974	0.788

Table 4. Average absolute difference, minimum (Min), maximum (Max) and average (Avr) MACE correlations for **Red Dairy Cattle** for protein, cell count, calving ease and female fertility (it) for correlation estimation using S-D deregressed files and S-MGS deregressed files as input for estimation of MACE correlations.

Trait	Avr abs difference	De-regressed SD-files			De-regressed S-MGS files		
		Min	Max	Avr	Min	Max	Avr
Protein	0.015	0.178	0.933	0.723	0.167	0.917	0.727
Cell count	0.016	0.686	0.963	0.873	0.683	0.963	0.879
Calving ease	0.008	0.890	0.983	0.943	0.903	0.999	0.948
Fertility (it)	0.079	0.012	0.907	0.557	-0.266	0.924	0.551

3.Results and Discussion

Changes in sire standard deviations (in percent) when changing from a Sire-MGS pedigree to a Sire-Dam pedigree for de-regression and sire variance estimation with data inclusion from 1970 onwards for protein, somatic cell count, direct calving ease and female fertility for HOL and RDC can be seen in Figure 1 and Figure 2, respectively. For most country-trait combinations an increase in sire standard deviation was observed. An increase in sire standard deviations when changing pedigree has earlier been reported by Van der Linde *et al.* (2005) and Jakobsen & Fikse (2009).

Average absolute difference, minimum, maximum and average MACE correlations for HOL and RDC for protein, cell count, calving ease and female fertility (it) for correlation estimation using S-D de-regressed files and S-MGS de-regressed files as input for estimation of MACE correlations are shown in Table 3 and Table 4, respectively. For protein yield and somatic cell count estimated correlations using either of the de-regressed files were very similar. This was also the case for female fertility (interval trait) for HOL and direct calving ease for RDC. For direct calving ease for HOL and female fertility for RDC correlations varied more with average absolute difference in correlations of 0.051 and 0.079, respectively. Deviations in correlations may be caused by a combination of low heritabilities, poor connetdness or that several national genetic evaluation centers are using Sire-MGS models for functional traits. Results show no evidence for an effect on correlations by using subsetting – no-subsetting for correlation estimation

In this study, national PGMs were de-regressed and sire variances estimated using data from 1970 onwards. For breeding value prediction in the Interbull routine evaluation only data from 1986 (HOL) and 1981 (other breeds) are used. Interest of international breeding values of bulls born before 1986 (1981) has increased due to a demand of MACE proofs for foreign bulls to be used in genomic evaluations. An option would therefore be to keep bulls back to 1970 also for the breeding value prediction but to re-estimate sire variances using the data reduced to 1986 (1981). This would alleviate use of the same de-regression for correlation estimation and breeding value prediction without changing current settled cut off years for sire variance estimation. And at the same time breeding values on older bulls could be computed. Different approaches for breeding value prediction are, however, left for further studies.

4. Conclusion

For protein yield, somatic cell count, female fertility (HOL) and direct calving ease (RDC) estimated correlations using either of the pedigree files were very similar. For direct calving ease (HOL) and female fertility (RDC) correlations varied more with average absolute difference in correlations of 0.051 and 0.079, respectively. Deviations in correlations may be caused by a combination of low heritabilities, poor connetdness or that several national genetic evaluation centers are using Sire-MGS models for functional traits.

5. References

De Jong, G. 2003. MACE – Options for improvement. *Interbull Bulletin 30,* 112-116.

Jakobsen, J. & Fikse, F. 2009. Sire-Dam pedigree in MACE – Results from a full scale pilot study. *Interbull Bulletin 39,* 91-97.

Klei, L. 1998. Solving MACE equations. Interbull Bulletin. *Interbull Bulletin 17,* 3-7.

Klei, L. & Weigel, K. 1998. A method to estimate correlations among traits in different countries using data on all bulls. *Interbull Bulletin 17,* 8-14.

Sigurdsson, A. & Banos, G. 1995. Dependent variables in international sire evaluations. *Acta Agric Scand., Section A, Anim. Sci. 45,* 209-217.

Sullivan, P. 1999. REML estimation of heterogeneous sire (co)variances for MACE. *Interbull Bulletin 22,* 146-148.

Van der Linde. C., de Roos, A.P.W., Harbers, A.G.F. & de Jong, G. 2005. Mace with sire-mgs and animal pedigree. *Interbull Bulletin 33,* 3–7.