

Implementing a Sire-Dam Pedigree Structure in MACE

J.H. Jakobsen and J.W. Dürr

Interbull Center, SLU, Box 7023, 750 07 Uppsala, Sweden

Abstract

The Interbull Community has decided to introduce a new MACE model including relationships on bull dams. The background for the change is to move genetic groups further away from animals with data in order for them to have less impact on the proofs. The consequences are changes in MACE breeding values (IEBV), especially for foreign bulls with no local progeny test, and an average increase in MACE reliabilities. The main reason for changes in IEBVs for this group of bulls is that the parent average of the bull is computed differently for sire-dam (SD)-MACE model compared to the sire-maternal-grandsire(S-MGS)-MACE model due to the change in the pedigree structure. The bull dam in the SD-MACE model gets a breeding value based on the relatives she has in the system. National breeding values (NEBV) of the bull dams are not included in the MACE model, and therefore the breeding value of the dam is only influenced by the performance of her relatives in the MACE system. A very positive performance of a dam, solely based on her relatives, will give a boost to a parent average of a bull compared to the parent average he had in the S-MGS system. A change in the parent average of a bull will therefore impact his converted IEBV to other country scales and is the main cause of changes between systems. The usage of the sire-dam pedigree in the MACE system contributes more information on the genetic background of a bull. In the case a dam has several sons tested in several countries; better links between countries are created.

Key words: international evaluation, MACE, SD-pedigree

Background

In every breeding value prediction, unknown parents are assigned to phantom parent groups (PHGs). These groups are of a certain size and in MACE based on animals of a certain origin, birth year and selection path. A group is treated as an animal and the same genetic group can have very different impact on the IEBV of the same animal in different country scales. In order to minimize the impact of the PHGs De Jong (2003) suggested including female relationships in the bull pedigree in MACE. This paper presents the main consequences of implementing a sire-dam pedigree in MACE and addresses the main concerns expressed by the Interbull customers during the pilots and test runs.

Changes to workflow

The only change to the MACE workflow is in the pedigree structure for de-regression and for breeding value predictions. For the correlation estimation part of the workflow the sire-maternal grandsire pedigree is still used for de-regression and correlation estimation. The changes are illustrated in Figure 1.

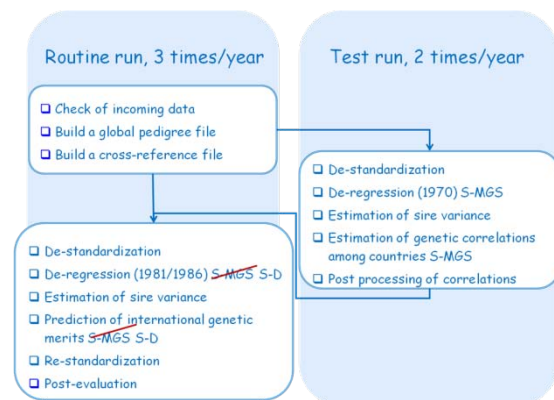


Figure 1. Changes to the MACE workflow.

Chronology

Following the idea by De Jong (2003) of adding bull dams to the pedigree it was decided to conduct a pilot run in order to investigate the impact of the change. Van der Linde *et al.* (2005) collected SD-pedigree information for seven Holstein populations and modified the software (Klei, 1998; Klei & Weigel, 1998) in order to run MACE with the modified pedigree structure. Van der Linde *et al.* (2005) found a better predictability of proofs but a 28 fold increase in CPU time with almost no change in correlations. Based on this the Interbull Technical Committee (ITC) decided to go for an implementation of breeding value prediction using SD-pedigree structure but to leave the correlation estimation as is.

Following the decision of an implementation of SD-MACE Interbull Centre started collecting more complete ancestry information on bull dams for all dairy breeds from all countries. Along with the collection of more complete pedigree also the software was further modified (Fikse, 2008).

The implementation by Van der Linde *et al.* (2005) was tested on protein BVs from seven well connected Holstein populations. Jakobsen & Fikse (2009) studied further if the SD-MACE model was applicable to traits with high and low heritabilities and populations with good and poor connectedness. The results looked promising and it was decided to go for a full scale implementation. With the continuous collections of pedigree information the limit was however reached for use of a direct solver and a solver based on iteration of data (Vouri *et al.*, 2006) was implemented along with a streamlining of the complete evaluation workflow.

The first official test run using sire-dam (SD) pedigree relationships was conducted in January 2011. After the results were released, concerns were raised about unexpected

decreases in reliabilities, increases in Mendelian Sampling variance $MS(\text{var})$ on foreign scales for French Red Holstein (FRR) bulls as well as the amount of re-ranking of bulls when compared to the S-MGS model. It was therefore decided not to implement the SD model in April 2011 and further investigate the outstanding issues.

Reliabilities are computed using the Harris and Johnson (1998) approach and are expected to increase by the increase in connectedness created by the inclusion of female relationships. Furthermore, the increase in $MS(\text{var})$ of FRR bulls on foreign scale was unexpected. Both issues were tailored to the new source of pedigree and were resolved by excess uploading of pedigree and especially linking IDs of animals pointing to the same animal. In cases where two aliases of an animal are treated as two individuals the information going into each of them is less than having the two aliases pointing to the same animal. The change in number of pedigree records and in number of cross references between January 2011 and January 2012 are shown in Table 1.

Table 1. Number of pedigree and cross (X) reference records in the Interbull pedigree database.

	January 2011	January 2012
Pedigree	4,166,833	4,333,257
X-reference	17,655	47,699

The Interbull pedigree database contains both dairy and beef but the increase in the number of pedigree records in the database between January 2011 and January 2012 are caused by a combination of actions in the dairy pedigree: inclusion of one new country (Korea) in the Interbull evaluations in September 2011; inclusion of pedigrees of young (unproven) bulls participating in the GMACE project; and upload of missing pedigree.

The cause of the re-ranking of bulls was tailored to the change in model by change in pedigree structure and was further investigated by looking into the bull dam pedigree path.

Bull dams in MACE model

National evaluation centers are often facing the problem of overestimated bull dams in national genetic evaluations. The difference in national evaluation models and international evaluation models is however that the proof of the dam is included in the national evaluation model while the dam only is included via pedigree relationships in the international model. MACE uses as input the breeding values of AI-tested bulls and the pedigree of the bulls. If the breeding value of a young AI-bull is overestimated in his national evaluation an overestimated breeding value will be submitted to Interbull. There is however no difference in national proof submitted to Interbull irrespectively if Interbull uses Sire-Maternal-Grandsire (S-MGS) or Sire-Dam (S-D) pedigree relationships.

The main difference of the S-MGS- and SD- MACE models is the addition of the bull dam into the pedigree relationships. This addition to the pedigree structure will link the sons she may have in the pedigree as well as link her to her own half sibs. Therefore, the addition of the bull dam will create a more dense pedigree structure on the female part of the pedigree. In order to investigate possible breed differences in number of progeny tested sons per bull dam, the percentage of the bull dams of different breeds of evaluations having one, two, three, four, five, six-to-ten, and more than ten sons is shown in Figure 2. The figure 2 illustrates that the Holstein breed of evaluation do have more bull dams with several AI sons compared to the other breeds.

Which bulls will have the largest changes in proofs

The average change in proofs by change of model was investigated for local and foreign bulls on the German scale. In this example, local bulls are bulls having a progeny test in Germany and foreign bulls are bulls with no progeny test in Germany. The amount of change in kg protein on German scale is illustrated in Figure 3. The figure shows that average changes are very small for local bulls but larger for foreign bulls.

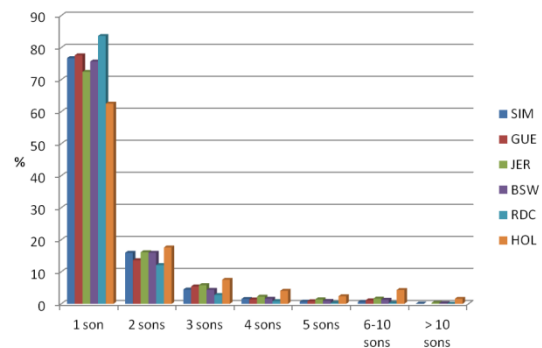


Figure 2. Percent of bull dams with one or more sons.

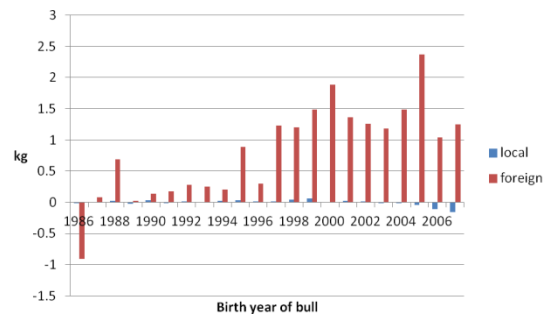


Figure 3. Average change in kg protein for Holstein proofs on German scale by change of model from S-MGS to SD-model.

A bull with a progeny test in only one country will get his proof converted to other country scales using Equation 1.

$$IEBV_{\text{bull}(\text{imp})} = PA_{\text{bull}(\text{imp})} + \text{cor}_{\text{imp,exp}} \left(\frac{SDV_{\text{imp}}}{SDV_{\text{exp}}} * (IEBV_{\text{exp}} - PA_{\text{bull}(\text{exp})}) \right) \quad [1]$$

where

$IEBV_{\text{bull}(\text{imp})}$ = International BV of bull in importing country

$PA_{\text{bull}(\text{imp})}$ = PA of bull in importing country

$\text{cor}_{\text{imp,exp}}$ = Genetic correlation between importing and exporting country

SDV_{imp} = Sire SDV in importing country

SDV_{exp} = Sire SDV in exporting country

$IEBV_{\text{exp}}$ = IEBV in exporting country

$PA_{\text{bull}(\text{exp})}$ = PA of bull in exporting country

Country of progeny test was named as exporting (exp) country and the country scale where the proof is converted to as importing (imp) country. When changing from a S-MGS pedigree structure to a SD-pedigree structure the sire standard deviations (SDV's) have also increased. An example of percentage change in sire SD when changing model can be seen in Figure 3 for parallel runs conducted during April 2011 and August 2011 for protein in Holsteins.

Both April and August are routine runs and no country is testing a new model in between these two runs. The changes in SDV's between the runs are therefore caused by updated national data and updated pedigree information in the Interbull pedigree database. For most countries the changes in SDV's were very similar for the two runs and all countries except Belgium had an increase in SDV by changing from the S-MGS- to the SD-pedigree. An increase in SDVs is expected due to more complete pedigree relationships (Sorensen & Kennedy, 1984).

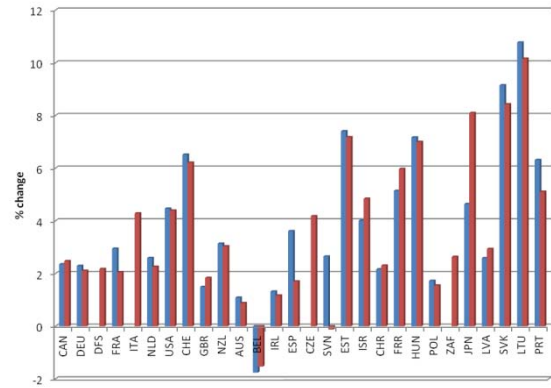


Figure 4. Percent changes in sire standard deviation for protein by change of model from S-MGS to SD-model. In red (April 2011) in blue (August 2011).

If sire SDs in all countries change with the same percentage, no re-ranking will occur due to these changes. However, what causes re-rankings due to changes in sire standard deviations is the ratio between two countries standard deviations. Therefore, if the increase in sire SD is the same for a two-country combination, the change in converted proofs is caused by factors other than the change in sire SD. Since the genetic correlations between countries are exactly the same in the two models, the only factor that can cause the change is the PA of the bull in the importing and in the exporting country. In the SD-MACE model parent average are computed according to the Equation 2:

$$PA_{\text{bull}} = \frac{1}{2} IEBV_{\text{sire}} + \frac{1}{2} IEBV_{\text{dam}} \quad [2]$$

where

PA_{bull} = International parent average of the bull

$IEBV_{\text{sire}}$ = International breeding value of the sire

$IEBV_{\text{dam}}$ = International breeding value of the dam

while parent average in the S-MGS model was computed according to Equation 3:

$$PA_{\text{bull}} = \frac{1}{2}IEBV_{\text{sire}} + \frac{1}{4}IEBV_{\text{mgs}} + \frac{1}{4}IEBV_{\text{mgd-gr}} \quad [3]$$

where

$IEBV_{\text{mgs}}$ = International breeding value of the maternal grand sire

$IEBV_{\text{mgd-gr}}$ = International breeding value for maternal grand dam group

As parent averages are computed differently than before, changes in PA's explain partially the observed differences in bulls IEBVs in foreign scales. The same does not occur in the local scale because the bull has a NEBV based on progeny test.

Bull dam families

The parent average of a bull in a SD-MACE model includes half the breeding value of the bull dam. The bull dam does not have a record of her own but does get a breeding value based on the performance of her relatives. If a bull dam has many relatives in the evaluation with a positive performance for that trait she will get a positive contribution from her relatives. On the other hand if many of her relatives in the evaluation have a negative performance for a trait she will also get a negative performance for the trait. These performances will in turn affect the PA of the bull. The PA of the bull is used in the conversion of a proof obtained in country of test to another country scale where the bull does not yet have a progeny test. Therefore, the bulls with the largest change in proof by change of model are the ones belonging to dam-families with very directional performances.

Conclusions

International proofs will change with change of pedigree structure. Proofs of foreign bulls will change more on local scale compared to proofs of local bulls on local scale. The main reason for the changes in proofs is due to inclusion of bull dam in the pedigree and along herewith the change in parent averages of the bull. Parent averages are used in the conversion of foreign bulls to local scale and do therefore impact the proofs.

Acknowledgements

Gerald Jansen is greatly acknowledged for streamlining of the SD-MACE software package currently used for international evaluations.

References

- de Jong, G. 2003. MACE – Options for improvements. *Interbull Bulletin* 30, 112-116.
- Fikse, F. 2008. Strategies to solve Mace mixed model equations. *Interbull Bulletin* 38, 146-151.
- Harris, B. & Johnson, D. 1998. Information source reliability method applied to Mace. *Interbull Bulletin* 17, 31-36.
- Jakobsen, J. & Fikse, F. 2009. Sire-dam pedigree in MACE - results from a full-scale pilot study. *Interbull Bulletin* 39, 91 – 98.
- Klei, B. 1998. Solving MACE equations. *Interbull Bulletin* 17, 3-7.
- Klei, B & Weigel, K. 1998. A method to estimate correlations among traits in different countries using data on all bulls. *Interbull Bulletin* 17, 8-14.

Sorensen, D. A. & Kennedy, B.W. 1984. Estimation of genetic variances from unselected and selected populations. *J. Dairy Sci.* 59, 1213-1223.

Van der Linde, R., 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.

Vuori, K., Strandén, I. & Lidauer, M.H. 2006. MiX99 – effective solver for large and complex linear mixed models. In: *Proc. 8th WCGALP, Belo Horizonte, Brazil*, CD-ROM, No 27: 33.