Illustration of an International Genetic Evaluation Robust to Inconsistencies of Genetic Trends in National Evaluations

Haifa Benhajali [(1)], **Jette Jakobsen** [(2)], **Sophie Mattalia**[(1)] **and Vincent Ducrocq**[(3)]
(1) Institut de l'Elevage,
(3) INRA,UMT 3G, INRA-GABI Bat 211, F 78352 JOUY en JOSAS cedex
(2) INTERBULL centre, Department of Animal Breeding and Genetics,
Swedish University of Agricultural Science, Box 7023, S-75007, Uppsala cedex
email: haifa.benhajali@idele.fr

Abstract

This paper presents a study of the consistency of MACE results in the Holstein population. Some biases are highlighted when a full pedigree structure is used in the MACE (SD MACE). They are probably due to discrepancies between genetic trends estimated by national evaluations and they were hidden when only male pedigrees were included (S-MGS MACE). These results do not question the advantages of a SD MACE in comparison with a S-MGS MACE, but some improvements are needed. Therefore a new model is proposed, in which the country effects are replaced by country x birth year effects. Results of a first test showed that this new model was able to correct international EBVs for the biases detected before. The Robust MACE would be an easy solution to obtain MACE results robust to over-or- under estimations of genetic trends, which are one of the main potential sources of bias in the national evaluations.

Introduction

MACE results are more and more important for national breeding programs, since they are now used not only to compare the genetic level of bulls available worldwide, but also as "pseudo-performances" of foreign bulls in reference populations used in genomic evaluations. Therefore, the maintenance of the quality of MACE is still strategic, even in the genomics era.

In 2012, the MACE model was changed by implementing a Sire-Dam pedigree structure in MACE (SD-MACE, Interbull, 2012) instead of a Sire – Maternal Grand Sire (S-MGS MACE). The main purpose of changing the pedigree structure was to move phantom parent groups of MGD further away from animals with data as it is well known that MGD-group solutions for a given bull can be rather different in different countries (De Jong, 2003). Group solutions are based on bulls tested in a country and linked to this group. Having phantom parent groups further away from animals with data will minimize their direct impact on international breeding values. This new model was therefore recognized by the Interbull community as an improvement, since it better takes selection of dams into account and because dams having sons in several countries create new links between countries. For most of the countries and for most of the traits, the changes due to the change in model were small. Some of them were relatively larger, such as in the case of production traits. For many years, selection in dairy cattle breeding has been mostly on production traits and change of the international pedigree structure was supposed to be able to also capture selection practiced on the bull dam path way. Thus a larger effect of the change was expected for traits under strong selection. However, some questions reminded and some deviations needed to be better understood. Surprisingly, the largest changes affected some bulls for longevity and for some type traits, with a variation in yearly average genetic levels reaching more than 0.2 genetic standard deviation.

The objective of this study was to analyze the consistency of MACE EBVs calculated with the new model. As it will be shown in the first section, some biases were identified, probably due to inconsistencies between genetic trends estimated in national evaluations. Therefore improvements in the MACE methodology are presented in the later sections, in order to get international comparisons robust to this source of bias.

In this paper, all the results are concentrated on progeny tested bulls from 4 different countries: France, the USA, Germany and the Netherlands. Many traits were analysed (production, type, fertility, longevity) on different scales. Only the two traits for which the change in model had the largest impact on the 4 countries, ie, Protein Yield and Longevity, are presented here.

Impact of the change in model on robustness against biases on genetic trend

In this section, the MACE EBVs used are those computed in December 2011 (last S-MGS MACE routine evaluation) and in April 2012 (first S-D MACE routine evaluation) in the Holstein population.

NB: In 2012, some countries had still incomplete pedigrees in the Interbull pedigree database, which could affect links between some dams and their own sire. However in the present study, we assumed that it did not affect the EBVs of the bulls of the 4 countries.

Two types of analyses were undertaken:

- *Mendelian sampling analysis*

On each country scale, Mendelian sampling (MS) was estimated for each bull as the difference between its MACE EBV and its parent average. In the S-MGS MACE, the parent average was computed by Interbull using the MACE EBVs of male ancestors and the estimated effect of the genetic group of maternal grand dams (MGD):

PA = ½ EBV Sire + ¼ EBV MGS + ¼ MGD Group.

In the SD-MACE, the EBV of the MGS and the MGD group effect are replaced by the MACE EBV of the dam:

PA = ½ EBV Sire + ½ EBV Dam.

Figures 1 to 4 show the average MS for bulls evaluated in the 4 countries, on the scale of country B. No clear trend in MS was detected with the S-MGS MACE model (for instance

0.01 kg/year for Bulls of country D for protein yield), while the MS averages of some countries with the new model depend on birth year with an average increase of 0.14 kg/year for the same bulls of country D). The same tendency is seen on other scales (for instance MS for protein yield for bulls of country D, expressed on the scale of country D, figure 5).

- *Full sibs analyses*

Families with at least 2 full sibs evaluated in 2 different countries were selected. The EBVs of their sons expressed in a given country-scale were analysed with the following model:

$$EBV_{ij} = Country_i + Family_j + e_{ij}$$

where Country i is the country with most of daughters and Family j is the jth family.

The analyses were done twice, first by taking all the full sibs into account, then by using only bulls born since 1995. In the latter situation, 1362 families of full sibs were considered for the test on Protein Yield, representing 874, 852, 722 and 584 US, German, French and Dutch bulls, respectively.

In absence of bias and as long as no pre-selection of bulls occurred before progeny test (which was still true for the evaluations from early 2012), estimated country effects were expected to be close to zero. This is the case when no selection on the age of the bulls is done (not shown here). When only bulls born since 1995 are selected, only differences between A and D and between A and B on protein yield are different from zero with the S-MGS MACE. With the SD MACE and for Protein Yield, the contrasts between A, B and C are not different from zero, while the results indicate an overestimation of full sibs of country D, when compared to those of countries A, B and C. With the SD MACE and for longevity, full sibs of country B seem to be underestimated, when compared to those of countries A, C and D. These results are consistent with the evolution of MS according to birth year: biases detected on young full sibs correspond to the countries for which a tendency on the average MS was detected, according to birth year. A positive trend (EBV> Pedigree index) reflects an overestimation of

the youngest bulls, a negative trend (EBV<Pedigree index) an underestimation.

In conclusion of these first two studies, some biases on MACE EBVs were suspected. They would depend on the age of the bulls and on their country of evaluation. They are much higher with the SD MACE than with the S-MGS MACE.

- *Left truncated evaluation*

In order to confirm these findings, Interbull ran an additional evaluation with data of April 2012 for longevity. In this run, the Sire Dam pedigree structure in MACE was maintained, but the EBVs of bulls born before 1995 were not taken into account, instead of in 1986, which is the of cut off year in conventional evaluations.

With this new evaluation, the non-zero country effects estimated with the full sibs analyses became non significant (tables 1 and 2). The estimated genetic trend was affected by truncation, leading for instance to a decrease in EBVs of foreign bulls expressed on country B – scale, while the introduction of the Sire – Dam pedigree structure had led to an increase of these EBVs (figure 6).

Discussion from the first studies, moving to a robust MACE

- Origin of the bias, role of the former MGD groups in case of inconsistencies between genetic trends

The first studies highlighted inconsistencies between SD MACE results, which were much less visible with the S-MGS MACE. Results of the "cut off 1995" evaluation show that **the comparisons of MACE EBVs of the most recent bulls are affected by the inclusion of national EBVs of old bulls.**

Moreover, since 2003, the oldest bulls have not been eliminated anymore from the international routine evaluations, as it was done each year before (and simulated in the "1995 truncated evaluation"). This means that the population of bulls with national EBVs included in MACE gets older and older. **Thus the**

problems that we observe today on the most recent bulls will probably be even more important in the future on the new bulls, if nothing is done.

The fact that these differences were much more visible with the SD MACE than with the S-MGS MACE is probably due to the fact that the national genetic trends are better captured with SD MACE than before. To understand better what is going on, we need to come back to the pedigree structure of the Holstein population. Interbull (2012) showed that most of the bull dams have few progenies. In most cases these dams have their progeny in their own country and dams used abroad are often of few other origins (mainly North American). In other words, with a S-MGS model, MGD groups are composed by dams of domestic bulls, except few MGD groups. Thus, step by step, Sire-MGS MACE forces the estimations of MGD group effects to be consistent with the genetic trend estimated by the national evaluation on the own country scale. For instance for a given country scale A:

- The estimated MGD group effects corresponding to the dams born in country A are mainly estimated from the national EBVs of the bulls evaluated in country A.
- If country A has many bulls born from dams of a given other country X, the groups corresponding to MGD of country X highly depend on the EBVs of the bulls evaluated in country A and having a dam born in country X. Thus, a potential bias in the national evaluation of country A also affects the estimated effects of the MGD groups of country X, when these effects are expressed on the scale of country A.
- The estimated genetic groups of MGD of country X must be consistent with those estimated for MGD of other origins. To illustrate this in a very simplistic case, if bulls evaluated in country Y and born from domestic dams or from dams of country X had a same sire and a same MGS and if they had a same national EBV, then the estimated genetic group effects of MGD of country Y and of country X would be the same.

With a SD MACE, no assumption is made on the genetic level of the MGD. The genetic trends estimated with SD MACE better

reflect the genetic trend estimated by the national evaluations, which is a good thing as long as the genetic trends estimated nationally are consistent. If not, the MGD groups of the previous model played a role of buffer, as they forced the pedigree indexes of foreign bulls to be consistent with the pedigree indexes of domestic bulls. The biases of the evaluations of foreign bulls have little impact on the EBVs of a given scale. And if the national genetic trend estimated in a given country scale is biased, the MGD group effects of all the origins are biased on the same way, with little impact on the comparisons of bulls of the same age.

- Presentation of the Robust MACE model

Ducrocq *et. al.* (2003) proposed an alternative approach relying on replacing deregressed EBVs by an annual DYD (= average DYD per year of performance and lactation number) of each sire. They showed that with a suitable sire model (including a year x country effect – instead of a single country effect), a lactation x country effect and a within country regression on age of the sire when his daughters are born), the genetic parameters and the sire solutions obtained were robust to over – or under estimation of genetic trend in national models. Obviously, the difficulty is to obtain annual DYDs from each participating country, particularly for more complex traits such as longevity or traits described by threshold models.

Nevertheless, this experience showed that replacing country effects by country x year effects had a positive impact on the results. Therefore, **a Robust MACE model is proposed** as a simplified version of the model of Ducrocq *et al.* (2003)**, in which country effects are replaced by fixed country x year effects, where year is the birth year of the bull.** When national genetic trends are consistent, the country effects are expected to be constant across birth years. If not, the country effects are expected to at least partly correct genetic effects for under – or over estimations.

Application of the Robust Mace model

- Data

Interbull ran a new evaluation with this Robust model (called "R-MACE" hereafter). In order to make all results comparable, the national deregressed EBVs, the within country sire variances and the genetic correlations were the same as in the April 2012 routine evaluation (SD MACE presented above).

Because of the small size of some populations, seven country-year per population were created, each country-year combination being defined by the bulls born within 3 successive years. As usual, all national proofs were de-standardized within country to a mean of zero and standard deviation of one before starting the process. After evaluation, the sum of EBVs + the country-year effect containing year 1996 scale were converted back to the original scale for each country.

Results

Consequences on genetic trends (figures 7 and 8)

The inclusion of country – birth year effects had consequences on the estimated genetic trends Even the genetic trend of the bulls evaluated on their own country scale could be affected by the change in model, as showed for country B for protein (slightly) and for longevity (high impact). This means that the national and the international EBVs may not be consistent anymore. In this case it should be recommended to publish MACE EBVs for all the bulls, even the domestic ones, as MACE is the only way to take worldwide selection into account.

Analysis of consistency of R MACE EBVs

Yearly average MS do not show any trend any more on Protein Yield (figure 9) and the range of within country variations is strongly reduced, when compared to the present SD MACE (now from -1.5 to +1.2 kg Protein with

the R-MACE against -1.8 to +2.3 with the official model). Similarly, there is no trend anymore on longevity (figure 10), but in this case the range of variations is not reduced for country B and variations seem to increase for the most recent years. This is probably due to a lack of information on the youngest bulls, and to the fact that our country-"year" effects merged 3 birth years. Moreover, in the full sibs analyses, the non zero country effects became non significant (table 2), except for longevity with country B. However, even in the latter case, these effects are reduced, compared to the present SD-MACE. This clearly shows the benefits of the use of the Robust MACE.

General conclusion

All the results of this study lead to the same conclusion: even if all the national evaluations passed the tests for genetic trend validations, some inconsistencies between estimated genetic trends still remain, and they affect MACE comparisons. SD MACE is better than S-MGS MACE because it moves phantom groups further way from animals with data, but at the same time it is probably more sensitive to any bias on national genetic trend.

Even if some country-traits are more affected than others, small discrepancies were identified in many cases. **In most cases, these small discrepancies do not question the global efficiency of MACE evaluations** (the best bulls, even if they are over estimated by 2 kgs of protein, are still excellent)**, but they may increase over time.** They affect the top rankings and the use of EBVs of young bulls, for instance in GMACE. Of course, each country should be encouraged to improve their national model of evaluation. But Ducrocq *et al.* (2003) showed that even if a national evaluation passes the validation tests, the genetic trend may be biased when a fraction of the bulls is selected based on foreign information, as the national evaluations only include national performances. Moreover, for MACE computations, it is recommended to send national EBVs computed with a single trait evaluation, even in the case of a trait highly correlated with some other highly selected traits. Therefore, some improvements must be found rapidly to obtain more robust international evaluations.

The first test on a (simplified) Robust MACE showed its ability to at least partly correct for discrepancies on national genetic trends. This simplified version also has the advantage to be easy to implement, since it does not require any new data.

At this stage, more studies are needed. In particular, it is suggested to separate country effects for each birth year when the population size is big enough. Also, the impact of a country x year effect on estimated genetic correlations should be studied, as genetic correlations are known to be biased downwards by discrepancies between genetic trends. Finally, the robustness of the model could be tested in a simulation in which a systematic bias could be included for one or many countries.

References

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

Ducrocq, V., Delaunay, I., Boichard, D. & Mattalia, S. 2003. A general approach for international genetic evaluations robust to inconsistencies of genetic trends in national evaluations. *Interbull Bulletin 30,* 101-111.

Interbull, 2012. Fact sheet about SD-MACE. *Available on www.interbull.org.*

Table 1. Contrast between country effects estimated for Protein Yield (unit scale = B) with several MACE models.

	Type of MACE		
	S-MGS	SD	Robust
A-D	1.86 ± 0.73*	-1.46 ± 0.69*	-0.31 ± 0.69
B-D	-0.66 ± 0.73	-2.32 ± 0.70**	-0.75 ± 0.70
C-D	0.37 ± 0.78	-2.18 ± 0.74**	-0.63 ± 0.75

(*: p<5%; **: p<1%)

Table 2. Contrast between country effects estimated for Longevity (unit scale = B) with several MACE models.

	Type of MACE			
	SMGS	SD	Cut off 1995	Robust
A-D	-0.02 ± 0.03	-0.04 ± 0.03	-0.03 ± 0.03	-0.05 ± 0.03
B-D	-0.03 ± 0.03	-0.18 ± .03**	-0.03 ± 0.03	-0.06 ± 0.03*
C-D	0.05 ± 0.03	-0.02 ± 0.03	0.02 ± 0.03	0.002 ± 0.03

(*: p<5%; **: p<1%)

Figures 1 to 4. Yearly average Mendelian samplings according to country of evaluation of the bulls, estimated from S-MGS MACE or S-D MACE (country scale = B).
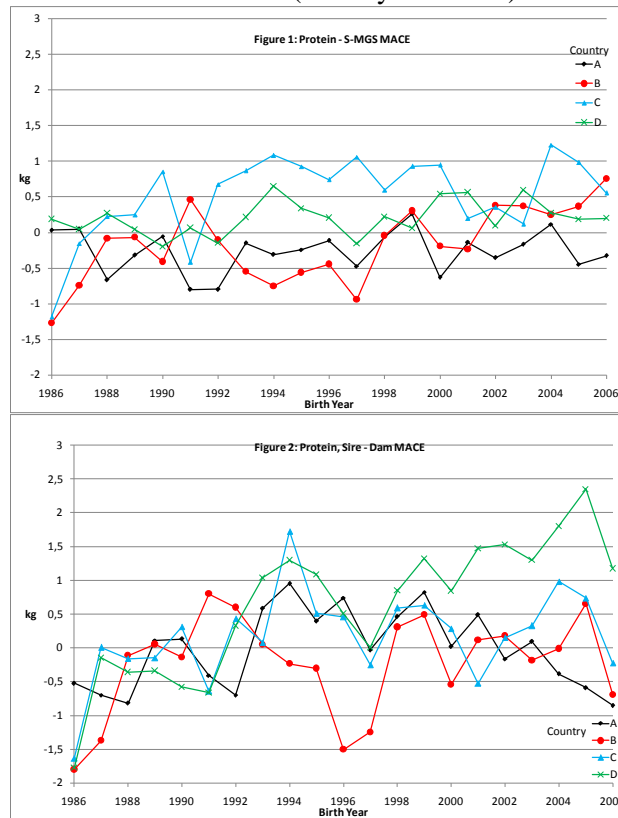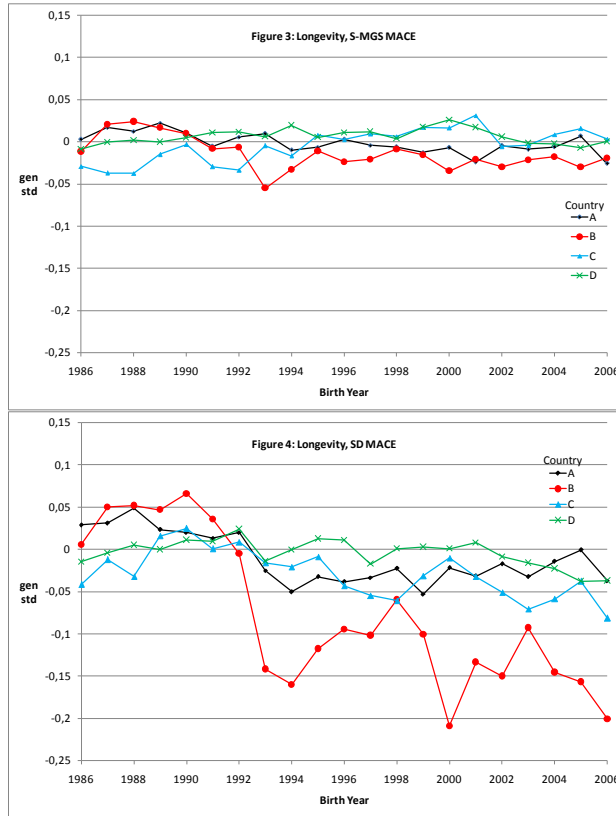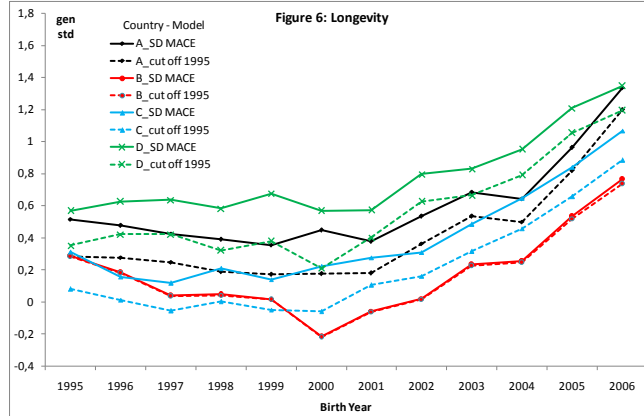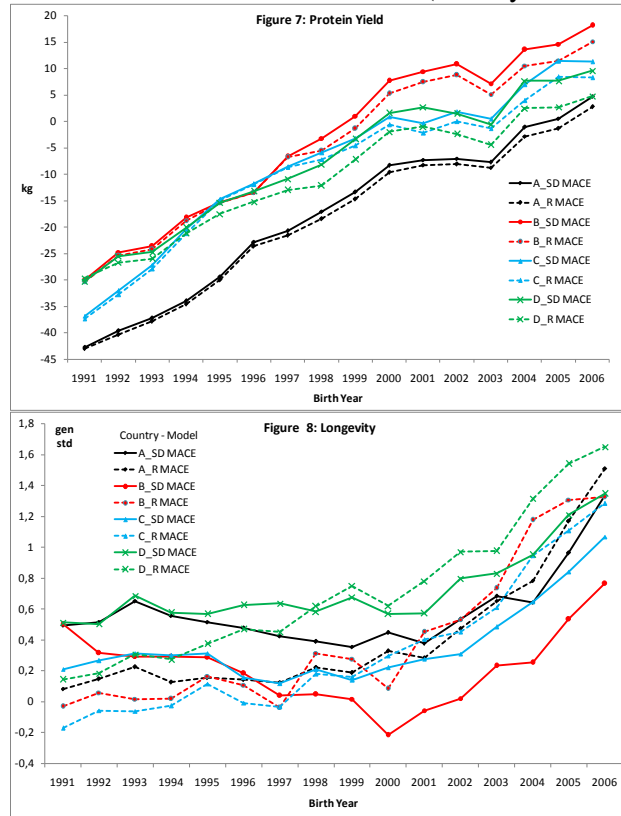
Figure 5. Yearly MS for protein Yield for bulls evaluated in country D, estimated from S-MGS MACE or SD-MACE (country scale = D).

Figure 6. Genetic trends of 4 countries on longevity, with the SD MACE or a cut off 95 SD MACE (country scale=B).

Figures 7 and 8. Genetic trends estimated for 4 countries, with SD MACE and Robust MACE (country scale=B).

Figures 9 and 10. Yearly average Mendelian samplings according to country of evaluation of the bulls, estimated from the Robust MACE (country scale=B).