GMACE Weighting Factors

P.G. Sullivan
Canadian Diary Network, Guelph, ON, Canada

Abstract

Statistical indicators to quantify information in national GEBV are required for optimal weighting of these GEBV as input observations for GMACE. The weighting factors are also used when approximating reliabilities of GMACE output, the international GEBV. The present study considered a change in approach from providing a recipe to national centers for deriving the weighting factors, to instead having Interbull apply the recipe internally and in a guaranteed consistent way for all countries. The main impacts of this change in approach were on selected estimates of genomic variance and GMACE reliabilities, most notably for specific estimates that had previously been questioned as erroneous. Most estimates of variance and reliability were essentially unaffected by the change in approach, because most weighting factors were only slightly different than before. An advantage of the new approach was that weighting factors and GMACE model specifics became perfectly aligned, and thus GMACE reliabilities were aligned without exception to the national values. All GMACE reliabilities were equal or higher than national values. All increases in reliability with GMACE were as expected, only for bulls that had input GEBV from more than one country, and with smaller increases if the multiple GEBV were from countries that share data for national genomic predictions.

Key words: genomics, international evaluation, GMACE, reliability

Introduction

Reliability approximations for GMACE were updated recently to better account for data contributions of maternal grand-sires and for performance information of dams (Sullivan, 2012a). These improvements aligned minimum reliabilities from GMACE with the national genomic reliabilities reported by all countries included in the present study. For many traits and countries, the GMACE and national genomic reliabilities are now very well aligned, but for some countries and traits, the GMACE reliabilities are generally higher than expected relative to national values. All contributions to GMACE reliability are via traditional (EDC) and genomic (GEDC) expected daughter contributions, which are the weighting factors for national genetic evaluations included as data in the GMACE model. The weighting factors are approximated by national evaluation centers following guidelines and methods approved by Interbull. The purpose of the present study was to investigate GMACE reliabilities that seem higher than expected, and to determine if weighting factors can be improved for GMACE.

Data

The data and edits for the present study were submitted to Interbull for the September 2012 test run of robust GMACE for young bulls. All traits in regular MACE were included in the test run, but the present study focused on only 5 of the traits; protein (pro), stature (sta), somatic cells (scs), direct longevity (dlo), and female fertility (cow conception trait #1; cc1. These are the same traits as in previous GMACE pilot studies (e.g. Jakobsen and Sullivan, 2012).

September 2012 national GEBV data from twelve populations (CAN, DEU, DFS, FRA, NLD, POL, USA, CHE, CHR, ITA, JPN and GBR), and EBV data from all participating countries in the September 2012 regular MACE test run of Interbull were used. The total numbers of official national GEBV on young genotyped bulls without daughter data, across

all populations, were: 83636 for pro, 90789 for sta, 79972 for scs, 60882 for dlo, and 59311 for cc1.

Methods

Weighting factors (GEDC) for young bull GEBV were computed by national evaluation centers as follows:

$$GEDC = k [GREL/(1-GREL) - REL/(1-REL)],$$

where k is the ratio of residual to genetic variance from the national genetic evaluation model, GREL is the national estimate of genomic reliability and REL the national estimate of reliability when ignoring genotypes. The GEDC and GREL were provided to Interbull as part of the data submission to participate in GMACE, but the REL used above were not included in the file formats. Determining REL by difference, it was apparent that definitions of REL were not consistent among all countries. For some countries, REL was 0 for all young bulls, while for other countries REL was generally in the range 0.25 to 0.40, which is a typical reliability range for parent averages.

An alternative weighting factor for GMACE, instead of the GEDC provided by countries, is the GEDC derived from the national GRELs provided relative to REL of MACE parent averages computed within the GMACE system. This alternative corresponds more directly with the evaluation variable (MS=GEBV-MACE PA) that is used for robust GMACE (rGM_ms(v)). While the MACE parent averages could be different from parent averages used nationally, it has become a common practice to include MACE proofs within national genomic evaluation systems. Parent average information contributes in a different way within genomic evaluation systems relative to pedigree-based evaluation systems, but GMACE uses pedigree-based evaluation methods to simplify and make feasible the extension of genomics to multiple countries (VanRaden and Sullivan, 2010).

In general, the use of parent average REL from MACE gave GEDC that were either similar or smaller than the GEDC provided, depending on the country. Lower GEDC decreased the reliabilities of Mendelian Sampling terms (rel(MS)) for some countries, and it became apparent that rel(MS) should be considered when computing a weighted average of the SDratios across countries. This weighted average is used to re-scale genomic SD estimates so that genomic evaluations are "MACE-neutral", i.e. comparable to evaluations from countries that do not have genomic systems in place yet. Weights for the average SDratio were thus changed from the number of young bulls to the sum of rel(MS), which matches the denominator term of the equation to estimate genomic variance (Sullivan, 2012b). This change had very minor effects on the estimates reported in the present paper, but improved adjustments for at least one of the other traits being studied as part of the September test run by Interbull.

An important advantage with the proposed approach of deriving GEDC from the national GREL relative to MACE REL, is the aligning of national and international genomic reliabilities. Bulls with a national GEBV that is official in only one country will have identical genomic reliabilities at the national and international levels, and bulls with multiple GEBV will have higher international relative to national genomic reliabilities. The amount that genomic reliabilities will increase depends on the amount of data sharing between the respective countries. If data sharing is high, the reliability increase will be small (Sullivan and VanRaden, 2010).

Results and Discussion

The effects of using different GEDC weighting factors on genomic variances are shown in Table 1, via ratios of genetic standard deviation estimates (genomic/MACE). These ratios tended to increase, by only small amounts but for all traits, in some countries and decrease in the other countries. Deriving GEDC from the national GREL instead of using the GEDC provided, which were derived separately by each country, was only different in the assumed value for REL of parent averages of the young bulls. The countries with increasing SDratios

had used smaller (or zero) REL than the REL of MACE parent averages when computing the GEDC to submit to Interbull.

Calculations of GEDC should be done consistently by all countries for GMACE, which would require additional communications, rules and/or guidelines, plus extra effort by all countries to eliminate the observed inconsistencies related to assumed values for REL. This extra work can be avoided if Interbull derives the GEDC from the GREL provided, as proposed in the present report. An additional advantage of using the national GREL is that GMACE and national GREL can be perfectly aligned, such that GREL from GMACE will always be equal or greater than the corresponding national GREL. This has been the general expectation of countries participating in GMACE, and deviations from this expectation were the main motivation for the present study.

Results in Table 2 demonstrate the deviations from expectation that have been observed for GMACE reliabilities when using the GEDC provided by the countries as weighting factors. For example, GREL changed with GMACE when no new information was available (i.e. bulls with GEBV in only 1 country). Results in Table 3 show that these deviations disappear when Interbull derives the GEDC weighting factors from the national GREL provided. If bulls have only 1 national GEBV included in GMACE, then the GREL from GMACE is always equal to the national GREL with this approach (Table 3). If bulls have multiple GEBV but only among countries within a single genomic data sharing group, then GREL increases with GMACE are relatively small, except where national GREL is much lower than for other countries (e.g. for GBR).

Increases are larger when bulls have GEBV from countries that are not part of the data sharing group, for example from DEU, NLD or FRA in addition to the 4 countries sharing data within the North American (NA) consortium, which includes USA, CAN, ITA and GBR. With the exception of bulls evaluated by GBR, GREL increased more by adding only 1 country outside the NA group than from all 3 of the other countries within the group. Countries in the NA group share almost all data on reference bulls. The Eurogenomics group, on the other hand, shares a smaller proportion of data from their respective reference populations, so GRELS would increase more than for NA bulls, if multiple Eurogenomics GEBVs were available for young bulls included in GMACE. All countries would benefit with higher GRELs if additional GEBVs were provided from any country.

Results in Table 4 confirm that GREL are completely aligned with national values when the national GREL are used to derive the GEDCs for GMACE. If GEDCs are computed by the national centers the GRELs are reasonably well aligned but there are many individual and country exceptions.

The Interbull file formats allow only 2 columns for GREL, with values ranging between 1 and 99. Allowing greater precision (e.g. 4 columns) could improve calculations of GEDC. However, for most traits and countries the 2-digit precision is adequate, because GMACE for young bulls limits the application to individuals with relatively large differences between REL and GREL. Greater precision for REL will become more important when extending the GMACE application to include bulls with daughters.

Acknowledgements

Appreciations extend to countries providing data for the September 2012 GMACE test run, and for helpful review and discussion by members of the GMACE working group; Jette Jakobsen, Gerben deJong, Zengting Liu and Paul VanRaden.

References

Jakobsen, J.H. & Sullivan, P.G. 2012. Robust GMACE for young bulls – application to data. *Interbull Bulletin 45*, 8-12.

Sullivan, P.G. and VanRaden, P.M. 2010. GMACE implementation. *Interbull Bulletin 41*, 3-7.

Sullivan, P.G. 2012a. GMACE reliability approximation. Report to the GMACE working group of Interbull. *Interbull Bulletin 37,* 1-4.

Sullivan, P.G. 2012b. GMACE variance estimation. Report to the GMACE working group of Interbull. *Interbull Bulletin 37,* 5-9.

VanRaden, P.M. and Sullivan, P.G. 2010. International genomic evaluation methods for dairy cattle. *Gen. Sel. Evol. 42, 7.*

Table 1. Ratio of genetic standard deviation estimates (genomic/MACE), with weighting factors derived from national GREL or based on the GEDC provided by national evaluation centers.

	Protein		Stature		SCS		Longevity		Fertility	
Country	GREL	GEDC	GREL	GEDC	GREL	GEDC	GREL	GEDC	GREL	GEDC
CAN	0.87	0.91	0.87	0.88	0.87	0.90	0.86	0.87	0.99	1.05
CHE	0.95	0.98	1.53*	1.60*	1.54*	1.58*	2.67*	2.38*	2.43*	1.89*
CHR	1.14	1.19	1.15	1.15	1.51*	1.40*	1.19	1.20	0.98	0.95
DEU	1.13	1.15	1.14	1.15	1.23*	1.26*	1.03	0.99	1.17	1.29*
DFS	1.17	1.09	1.23*	1.07	1.31*	1.26*	-	-	1.43*	1.44*
FRA	0.92	0.97	1.03	1.12	1.06	1.14	0.94	1.05	0.92	1.05
GBR	1.11	1.04	1.00	0.99	0.89	0.86	0.79*	0.71*	0.79*	0.72*
ITA	1.03	1.01	0.97	0.94	0.99	0.97	1.25*	1.18	-	-
JAP	1.06	1.02	1.14	1.07	1.19	1.09	-	-	-	-
NLD	1.12	1.04	1.13	1.09	1.17	1.08	0.67*	0.55*	0.81	0.61*
POL	1.19	1.13	1.18	1.10	1.13	1.07	-	-	1.19	1.07
USA	0.96	0.96	1.02	1.01	1.00	1.00	1.05	1.03	1.10	1.14

*Estimates that are truncated to the limits of range [0.80,1.20].

Table 2. National Genomic reliabilities (GREL) for protein, and GREL increases with GMACE, for bulls with GEBV from different combinations of countries. Weighting factors in GMACE were GEDC derived and provided by the national evaluation centers.

Country	National GREL		Average Increase in GREL with GMACE					
			GEBV in 1 country		*NA GEBV (4 countries)	NA+DEU (5 countries)	NA+FRA (5 countries)	NA+NLD (5 countries)
	n	Ave	n		14855	98	83	26
CAN	18648	73.3	230	0.0	4.5	7.6	7.2	8.5
GBR	17641	63.1	880	5.6	11.9	14.2	14.1	13.7
ITA	17685	74.2	299	2.8	4.5	6.6	6.3	5.8
USA	18735	76.5	474	3.5	3.7	5.7	5.3	5.3
DEU	2350	73.3	1980	0.5		7.5		
FRA	1058	71.0	819	-0.3			9.2	
NLD	3825	62.0	3762	5.9				17.3

*North American data sharing consortium, including Canada and USA, plus Italy and Great Britain

Table 3. National Genomic reliabilities (GREL) for protein, and GREL increases with GMACE, for bulls with GEBV from different combinations of countries. Weighting factors in GMACE were GEDC derived by Interbull from GREL provided by the national evaluation centers.

Country	National GREL		Average Increase in GREL with GMACE					
			GEBV in 1 country		*NA GEBV (4 countries)	NA+DEU (5 countries)	NA+FRA (5 countries)	NA+NLD (5 countries)
	n	Ave	n		14855	98	83	26
CAN	18648	73.3	230	0.0	3.1	6.6	6.0	6.7
GBR	17641	63.1	880	0.0	8.4	11.5	11.4	10.3
ITA	17685	74.2	299	0.0	1.7	4.4	4.0	3.2
USA	18735	76.5	474	0.0	1.0	3.9	3.4	2.9
DEU	2350	73.3	1980	0.0		6.5		
FRA	1058	71.0	819	0.0			8.0	
NLD	3825	62.0	3762	0.0				14.2

*North American data sharing consortium, including Canada and USA, plus Italy and Great Britain

Table 4. Overall distributions of national GREL and changes with GMACE based on two different sets of GEDC weighting factors.

		National GREL	GREL change with GMACE using GEDC derived nationally			GREL change with GMACE using GEDC derived by Interbull		
Country	n	average	average	minimum	maximum	average	minimum	maximum
CAN	18648	73.3	4.4	-4	30	3.1	0.0	27
CHE	888	75.0	2.4	-1	9	2.3	0.0	8
CHR	927	66.9	3.5	-1	17	3.7	0.0	16
DEU	2350	73.3	1.5	0	10	0.9	0.0	9
DFS	865	66.1	6.4	5	16	0.0	0.0	11
FRA	1058	71.0	1.8	-1	13	1.8	0.0	12
GBR	17641	63.1	11.6	3	18	8.0	0.0	16
ITA	17685	74.2	4.4	1	14	1.6	0.0	13
JPN	1014	66.3	4.7	2	15	0.0	0.0	0
NLD	3825	62.0	6.1	4	21	1.5	0.0	19
USA	18735	76.5	3.7	0	18	1.1	0.0	14