Defining a Parameter Space for GMACE

P.G. Sullivan
Canadian Dairy Network, 660 Speedvale Avenue West, Suite 102, Guelph, Ontario, Canada N1K 1E5

Abstract

National genomic evaluations of young bulls (GEBV) are combined by Interbull, using an international genomic MACE model (GMACE), with non-zero residual correlations to account for sharing of genotypes among national genomic evaluation systems. It was observed recently that GMACE results for mastitis resistance were inconsistent with corresponding results for somatic cell score. This study examined the current GMACE methods, and new modifications to better account for different heritabilities and for different genomic reliabilities among countries for a given trait. A parameter space was defined that bounds GMACE results, on the scale of each country, to fall somewhere between the national GEBV, and predictions of international GEBV when sharing of genotypes, common SNP panels, etc, are ignored. Distances to either boundary were estimated as a function of the degree of data sharing observed among national genomic evaluation systems. The proposed modifications to GMACE had largest effects on traits with a wide range of heritabilities among countries, such as mastitis resistance, and on the scales of countries that had relatively low national genomic reliabilities. Results from GMACE were much more consistent between mastitis and somatic cell score after the modifications, and also among all other traits evaluated by Interbull.

Key words: genomics, reliability, international evaluation, GMACE, parameter space

Introduction

Somatic Cell Score (SCS) is a genetically correlated trait used to indirectly select for improved mastitis (MAS) resistance. Both the heritabilities and volumes of data are much higher, and as a consequence reliabilities are higher for SCS than for clinically observed MAS, in countries that evaluate both traits.

Interbull provides international evaluations for both of these traits. For MAS, national evaluations for MAS are used if available, but otherwise the national evaluations for SCS are used as a predictor trait. Only SCS evaluations are used in the international evaluations of SCS. The international reliabilities for MAS should rarely if ever be higher than the corresponding international reliabilities for SCS. While the MACE reliabilities for MAS and SCS follow this expected pattern, some of the GMACE reliabilities have been higher than expected for MAS (G. deJong, pers. comm., 2015).

Objectives of the present study were to find explanations for relatively high GMACE reliabilities for MAS, and to improve the GMACE evaluation system, if possible.

Materials and Methods

The input data and published results, from December 2015 GMACE evaluations of MAS and SCS, were reviewed to determine if and why MAS reliabilities were too high. While residual correlations have a direct impact on reliabilities, these parameters are not estimated from the phenotypic data used in GMACE, and were of particular interest in this study.

Simulated data were used to study patterns of GMACE reliabilities under different input data scenarios, to compare alternative assumptions for residual correlations and to study new modeling approaches for GMACE. Data were simulated for an individual animal,

with national genomic evaluations from each of two groups of countries sharing genomic data (genomic consortia) to increase genomic evaluation reliabilities at the national level. Each consortium included 4 countries. The degree of data sharing was 0% between consortia and was varied from 0 to 100% within consortia. Genetic correlations among countries were 0.80 between consortia and 0.90 within consortia. National genomic reliabilities and trait heritabilities were varied, where countries with lower trait heritabilities also had lower genomic reliabilities. These data were similar to the simulated data used to develop GMACE systems currently used by Interbull (Sullivan and VanRaden, 2010), but with the added complexity of varying trait heritabilities among the countries.

Results from the simulated data were used to identify and test new GMACE methods that take better account of variation among countries in trait heritabilities, and also among genomic reliabilities (Appendix). The new GMACE methods were then tested on more recent data, and for all 38 traits included in the September 2016 Interbull GMACE test run.

Results & Discussion

December 2015 Data

The residual correlations in the official GMACE evaluation (GMACE model 1) were generally much lower for MAS than for SCS. Below is an example to demonstrate, for a bull with genomic evaluations from both the North American and EuroGenomics consortia. Although no national reliabilities were higher for MAS than SCS for this bull, the GMACE reliabilities were higher for MAS on 5 of these 8 country scales (ITA, GBR, NLD, ESP and DEU). Related to the higher MAS reliabilities, for this and other bulls, were correspondingly higher variances of GMACE EBV (results not shown), consistent with the theoretically proportional relationship between reliability and EBV variance. When GMACE reliabilities are biased upwards, the variance of GMACE EBV are also biased upwards to a similar degree.

Table 1. December 2015 published reliabilities and GMACE correlations assumed for bull X.

Eval'n Scale[z]	National Reliability		GMACE Reliability		Correlations ($\bar{r}_e / \bar{r}_g$)%	
	SCS	MAS	SCS	MAS	SCS	MAS
CAN	70	33	81	61	27	9
ITA	53	-	76	83	19	5
GBR	67	-	80	84	32	9
FRA	68	56	80	72	40	28
NLD	67	64	78	81	25	17
DFS	73	68	82	79	36	21
ESP	67	-	78	85	38	13
DEU	75	-	82	88	38	13

[z]Countries listed in Table 10.

New residual correlations (GMACE model 2), and GMACE reliabilities, were derived for the same bull (Table 2), with an adjustment to account for different heritabilities among country-traits (Appendix). The new residual correlations were generally higher than before, and were now much more similar between MAS and SCS. The patterns of difference between GMACE and national reliabilities also became more similar and reasonable, for MAS relative to SCS, with the MAS reliabilities more generally equal or lower than SCS on all national scales of evaluation. One small remaining exception was for country ITA.

Table 2. December 2015 GMACE reliabilities using new residual correlations, for bull X.

Eval'n Scale[z]	National Reliability		GMACE Reliability		Correlations ($\bar{r}_e / \bar{r}_g$)%	
	SCS	MAS	SCS	MAS	SCS	MAS
CAN	70	33	78	56	32	19
ITA	53	-	74	75	23	21
GBR	67	-	78	77	31	27
FRA	68	56	78	68	47	39
NLD	67	64	75	73	49	44
DFS	73	68	79	74	53	46
ESP	67	-	76	76	49	46
DEU	75	-	80	80	56	51

[z]Countries listed in Table 10.

Reliability increases were notable but relatively small for countries with high national reliabilities, and were much larger for countries with low national reliabilities. Additional changes (Appendix) were made to constrain GMACE reliabilities and EBV variance on all scales of evaluation, based on a parameter space that especially affected the scales with low national reliabilities (GMACE model 3). Results are in Table 3, again for the same bull. With the parameter space restrictions, all MAS reliabilities for this bull were consistently equal or lower than SCS reliabilities, with no exceptions.

Table 3. December 2015 GMACE reliabilities constrained to a parameter space, for bull X.

Eval'n Scale[z]	National Reliability		GMACE Reliability		Correlations ($\bar{r}_e / \bar{r}_g$)%	
	SCS	MAS	SCS	MAS	SCS	MAS
CAN	70	33	76	53	32	12
ITA	53	-	67	67	23	19
GBR	67	-	76	75	30	24
FRA	68	56	76	64	46	32
NLD	67	64	75	71	45	38
DFS	73	68	78	72	53	42
ESP	67	-	75	75	46	41
DEU	75	-	80	79	55	48

[z]Countries listed in Table 10.

Simulated Data

The three GMACE models described above were also considered in this section:

1) Interbull application in December 2015.
2) Modified to account for trait heritability.
3) Using parameter space restrictions.

Differences between Models 2 versus 1 were relatively small. Maximum GMACE reliability was reduced to exactly match the national reliability with Model 2, when all data were shared among national genomic evaluations (100% shared reference populations; Tables 4 and 5). However, for countries with lower national reliabilities, the model 2 reliabilities were higher than model 1. The maximum reliability was identical between models 3 and 2, while model 3 reliabilities were always lowest for countries with lowest national reliabilities. Differences between GMACE and national reliabilities were smallest, as expected, with model 3.

Table 4. A bull with GEBV from 4 countries sharing genotypes of genomic reference bulls.

Country		A	B	C	D
Trait Heritability		11	20	20	33
National Reliability		41	49	60	70
% Shared Reference	GMACE Model	GMACE Reliability			
100	1	62	62	65	72
	2	63	63	65	70
	3	41	49	60	70
50	1	64	65	67	71
	2	66	67	69	72
	3	57	61	67	72
0	ALL	72	73	75	77

Results in Table 5 confirmed that GMACE model 3 avoids double-counting the shared national genomic information, while models 1 and 2 do not completely avoid the double-counting. Each of the 4 countries in Table 4 was duplicated (4 pairs of countries) in Table 5. With 100% shared reference populations, reliabilities from model 3 were the only ones unaffected by the duplicating of countries (i.e. identical reliabilities in Tables 4 and 5).

Table 5. A bull with GEBV from 8 countries sharing genotypes of genomic reference bulls.

Country(consortium)		A_1 E_1	B_1 F_1	C_1 G_1	D_1 H_1
Trait Heritability		11	20	20	33
National Reliability		41	49	60	70
% Shared Reference	GMACE Model	GMACE Reliability			
100	1	65	66	67	74
	2	64	64	66	70
	3	41	49	60	70
50	1	64	65	67	71
	2	66	67	69	72
	3	60	64	69	74
0	ALL	79	80	81	83

Results in Table 6 are more representative of current data evaluated by Interbull, where data sharing is generally high but below 100%

within (e.g. 90%) and low between (e.g. 25%) two genomic consortia. These patterns of reliability, for 90 : 25 % Shared Reference, indicate how official GMACE reliabilities can be expected to change when switching from models 1 to 3, for a bull genomically evaluated in both North America and EuroGenomics.

Table 6. A bull with GEBV from 8 countries, divided between 2 separate consortia.

Country(consortium)		A_1 E_2	B_1 F_2	C_1 G_2	D_1 H_2
Trait Heritability		11	20	20	33
National Reliability		41	49	60	70
% Shared Reference[z]	GMACE Model	GMACE Reliability			
100 : 0	1	68	69	71	77
	2	69	70	71	75
	3	61	64	69	75
90 : 25	1	67	67	69	75
	2	68	68	70	73
	3	57	61	70	73
50 : 0	1	71	72	73	76
	2	72	73	75	77
	3	69	70	73	77
0 : 0	ALL	76	77	79	81

[z]within : between consortia.

September 2016 Data

Model 3 was applied to the GMACE test run data for all traits, to directly compare results against the model 1 test run results distributed to all countries by Interbull.

Changes were only expected for bulls having national genomic evaluations in more than 1 country. It was confirmed that for bulls with only 1 national GEBV, the model 3 GMACE results were identical to model 1, and in each case matching the national GEBV and reliabilities provided. Among the bulls with national GEBV from multiple countries, there are two main groups; those evaluated in either 1 or in both of the North American and EuroGenomics consortia. GMACE results for bulls evaluated in only one consortium are expected to be very consistent with the national results, due to very high levels of data sharing within each consortium. More significant changes are expected for bulls evaluated in both consortia, because sharing of genomic data is low between consortia.

Results in Table 7 are for the bulls evaluated in only 1 consortium. Changes from national results were much smaller with GMACE model 3 compared with model 1, across all traits. For these bulls, the average increase over national reliability was about 1 point with GMACE model 3, and the change in GEBV variance was very close to zero.

For bulls evaluated in both consortia, changes from national results were reduced with model 3, relative to model 1 (Table 8), but were still substantial. Reliabilities increased from national values by about 7 points with GMACE model 3, and there was a similar increase in variance of GEBV. Genomic selection of bulls in this group can be improved by using GMACE rather than national genomic evaluations, under the current scenario where sharing of reference population genotypes is minimal between the North American and EuroGenomics consortia.

Bias of genomic rankings with GMACE was assessed by the results in Table 9. Under GMACE models 1 and 3, regressions of GMACE on national GEBV were generally close to the expected value of 1.0 across all traits and evaluation scales. Differences from expected regression of 1.0 were generally closer to zero under model 3. For example, the standard deviation of this difference (observed - expected $\hat{b}$), across all traits and scales of evaluation, was reduced by 50% for the bulls with GEBV in 1 consortium (from 0.036 to 0.018), and from 0.064 to 0.059 for bulls evaluated in both consortia (Table 9).

Table 7. GMACE relative to national results by trait, averaged across scales of evaluation, for bulls evaluated in only 1 consortium (standard deviation in parentheses).

	Average Increase in Reliability		Average Ratio of σ(GEBV) - 100%	
	GMACE Model		GMACE Model	
Trait[z]	1	3	1	3
ANG	3.6	0.6	3.0	0.5
BCS	3.6	0.7	0.7	-1.5
BDE	3.0	1.2	1.0	0.6
CC1	2.7	0.7	1.9	0.4
CC2	4.9	1.3	5.6	-0.1
CRC	3.8	1.3	4.1	1.8
CWI	2.9	1.1	2.1	0.8
DCE	1.5	0.2	3.7	0.5
DLO	2.9	0.3	4.3	0.3
DSB	1.2	0.1	0.8	-0.2
FAN	3.5	0.9	1.8	0.3
FAT	3.2	1.1	1.9	-0.6
FTL	4.0	1.6	1.6	0.1
FTP	3.7	1.6	2.0	0.6
FUA	2.2	1.1	-0.2	0.4
INT	4.3	1.5	3.4	0.0
LOC	3.2	0.4	2.5	0.1
MAS	5.8	1.1	9.9	0.9
MCE	1.7	0.2	-0.7	-0.2
MIL	3.4	1.2	1.3	-1.3
MSB	2.5	0.2	1.2	0.0
MSP	1.6	0.4	0.0	-0.5
OCS	1.8	0.4	-0.3	-0.7
OFL	2.9	0.6	2.4	0.3
OUS	1.5	0.6	0.3	-0.1
PRO	2.8	1.0	1.0	-0.9
RAN	3.9	1.8	1.0	-0.1
RLR	3.4	1.1	1.4	0.3
RLS	3.3	1.3	2.6	1.1
RTP	4.2	1.8	1.0	0.4
RUH	2.8	1.2	-0.3	-0.5
RWI	3.1	1.5	0.9	0.9
SCS	3.7	1.3	3.1	1.3
STA	4.5	2.0	0.3	-0.6
UDE	3.8	1.5	0.4	0.1
USU	2.7	1.1	1.9	0.4
Average	3.2 (1.0)	1.0 (0.5)	1.9 (2.0)	0.1 (0.7)
All Traits/ Scales	3.2 (2.4)	1.0 (1.0)	1.9 (4.5)	0.1 (1.9)

[z]Traits are described in Table 10.

Table 8. GMACE relative to national results by trait, averaged across scales of evaluation, for bulls evaluated in 2 consortia (standard deviation in parentheses).

	Average Increase in Reliability		Average Ratio of σ(GEBV) - 100%	
	GMACE Model		GMACE Model	
Trait[z]	1	3	1	3
ANG	8.6	5.3	9.3	3.8
BCS	9.9	9.8	4.3	3.0
BDE	8.9	6.7	5.1	3.9
CC1	8.3	5.2	5.9	3.3
CC2	12.8	7.3	13.4	4.7
CRC	10.5	7.8	8.6	5.0
CWI	8.6	6.8	5.8	2.6
DCE	7.2	5.0	8.2	4.2
DLO	8.4	6.4	10.8	4.0
DSB	8.1	4.7	7.6	0.4
FAN	8.9	7.0	5.3	1.4
FAT	8.2	5.7	5.4	1.9
FTL	10.3	7.6	5.4	2.5
FTP	9.7	7.5	4.8	2.1
FUA	7.4	6.4	2.9	2.9
INT	11.0	7.5	10.1	4.3
LOC	9.1	6.5	7.3	3.5
MAS	13.8	7.8	18.1	9.4
MCE	6.5	5.0	2.3	1.7
MIL	8.2	5.9	4.3	0.4
MSB	7.7	5.4	2.4	0.4
MSP	7.5	6.1	6.0	4.0
OCS	7.7	5.6	3.4	0.9
OFL	8.2	6.5	4.9	1.4
OUS	6.5	5.5	2.8	1.8
PRO	7.7	5.6	3.6	0.4
RAN	10.3	8.2	4.4	2.9
RLR	10.0	7.0	8.3	2.6
RLS	9.0	7.4	5.1	3.5
RTP	10.5	8.7	5.5	3.9
RUH	8.1	6.9	3.9	2.9
RWI	9.1	7.3	4.8	3.0
SCS	9.5	7.0	8.5	5.3
STA	10.1	8.1	2.4	0.3
UDE	9.7	7.6	4.4	3.4
USU	8.0	6.2	5.9	3.0
Average	9.0 (1.5)	6.7 (1.2)	6.1 (3.3)	2.9 (1.8)
All Traits/ Scales	9.0 (3.0)	6.7 (2.8)	6.2 (8.8)	2.9 (7.6)

[z]Traits are described in Table 10.

Table 9. Observed regressions of GMACE on national GEBV by trait, relative to expectation (100%*($\hat{b}$-1)), averaged across scales of evaluation (standard deviation in parentheses).

	GEBV in only 1 Consortium		GEBV in 2 Consortia	
	GMACE Model		GMACE Model	
Trait[z]	1	3	1	3
ANG	1.6	0.3	5.3	1.4
BCS	-4.2	-2.2	-5.3	-6.0
BDE	0.3	0.1	1.3	1.1
CC1	-0.9	-0.1	-1.9	-1.3
CC2	3.1	-0.7	5.7	0.0
CRC	2.5	1.1	2.5	0.4
CWI	0.9	0.4	1.4	-0.5
DCE	2.3	0.3	2.7	0.5
DLO	2.2	-0.1	4.5	0.4
DSB	-0.4	-0.2	0.6	-3.4
FAN	0.1	0.0	-0.1	-2.1
FAT	1.0	-0.9	2.6	-0.5
FTL	0.5	-0.4	1.5	-0.8
FTP	0.5	-0.3	0.6	-1.5
FUA	-0.8	-0.1	-0.5	0.1
INT	1.4	-0.7	2.1	-0.9
LOC	-0.5	-0.4	-1.4	-2.9
MAS	5.8	-0.3	6.3	0.9
MCE	-1.7	-0.2	-1.7	-1.7
MIL	0.0	-1.9	1.0	-2.0
MSB	-0.4	-0.2	-1.6	-2.6
MSP	-1.5	-0.8	0.2	-1.0
OCS	-1.4	-0.9	-0.9	-2.4
OFL	0.1	-0.3	-1.1	-3.0
OUS	-0.1	-0.4	-0.5	-0.8
PRO	0.0	-1.5	0.8	-1.9
RAN	-0.5	-0.5	0.0	-0.9
RLR	-1.1	-0.4	0.9	-2.4
RLS	1.4	0.6	1.4	0.3
RTP	-0.5	-0.5	1.0	0.3
RUH	-1.4	-0.9	-0.1	-0.4
RWI	-0.3	0.3	1.1	0.0
SCS	1.9	0.5	3.8	1.4
STA	-1.0	-1.5	-0.9	-2.6
UDE	-1.3	-0.8	0.1	-0.4
USU	1.0	0.1	2.2	0.3
Average	0.2 (1.7)	-0.4 (0.7)	0.9 (2.3)	-1.0 (1.6)
All Traits/ Scales	0.3 (3.6)	-0.4 (1.8)	1.0 (6.4)	-0.9 (5.9)

[z]Traits are described in Table 10.

Table 10. Codes used by Interbull.

Evaluation Scale	Country
CAN	Canada
ITA	Italy
GBR	Great Britain
FRA	France
NLD	Netherlands
DFS	Denmark-Finland-Sweden
ESP	Spain
DEU	Germany
Trait	Description
ANG	Angularity
BCS	Body Condition Score
BDE	Body Depth
CC1	Cow Conception Trait 1
CC2	Cow Conception Trait 2
CRC	Cow Return to Conception
CWI	Chest Width
DCE	Direct Calving Ease
DLO	Direct Longevity
DSB	Direct Stillbirth
FAN	Foot Angle
FAT	Fat Yield in the Milk
FTL	Front Teat Length
FTP	Front Teat Placement
FUA	Fore Udder Attachment
INT	Fertility Interval Trait
LOC	Locomotion
MAS	Mastitis Resistance
MCE	Maternal Calving Ease
MIL	Milk Yield
MSB	Maternal Stillbirth
MSP	Milking Speed
OCS	Overall Conformation Score
OFL	Overall Feet and Legs Score
OUS	Overall Udder Score
PRO	Protein Yield in the Milk
RAN	Rump Angle
RLR	Rear Legs Rear View
RLS	Rear Legs Side View
RTP	Rear Teat Placement
RUH	Rear Udder Height
RWI	Rump Width
SCS	Somatic Cell Score
STA	Stature
UDE	Udder Depth
USU	Medial Udder Support

Conclusions

Evidence from the present study supports the conclusion that GMACE reliabilities were inflated for MAS, and to a lesser degree many other traits, due to a double-counting of shared genomic data at both the national and international levels. New GMACE methods were developed to better avoid double-counting the shared genomic data, more consistently across all traits and scales of evaluation. The new methods will be implemented by Interbull in January 2017.

Acknowledgements

Detailed reviews of official GMACE results by Gerben deJong, and subsequent discussions with Hossein Jorjani, were very helpful for designing this study. Co-ordination of data and server access at the Interbull centre was much appreciated, as provided by Haifa Benhajali and Carl Wasserman.

References

Sullivan, P.G. & Jakobsen, J.H. 2012. Robust GMACE for young bulls – methodology. *Interbull bulletin 45*, 3-7.

Sullivan, P.G. & Jakobsen, J.H. 2014a. GMACE Pilot #4: Adjusting the National Reliability Input Data. *Interbull Bulletin 48*, 40-45.

Sullivan, P.G. & Jakobsen, J.H. 2014b. GMACE without variance estimation. *Interbull Bulletin 48*, 46-49.

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

Appendix

If genomic data sharing was 100% among all countries, then all output reliabilities should equal the corresponding input values, because all available genomic data were already used at the national level. To satisfy this condition, a residual correlation for GMACE can be derived for a 2-country example, using selection index to combine averages of δ_1 and δ_2 effective records, respectively.

Let $x_i = 1 + \alpha_i / \delta_i$, where $\alpha_i = R_{ii} / G_{ii}$. The x_i is inverse of reliability given δ_i, $V_{ii} = x_i G_{ii}$ and $V_{ij} = wG_{ij}$. We need to solve for *w*, such that output reliability is equal to input reliability:

$$(\mathbf{GV^{-1}G})_{ii} / G_{ii} = (G_{ii} V_{ii}^{-1} G_{ii}) / G_{ii}$$
$$(\mathbf{GV^{-1}G})_{ii} = (G_{ii} V_{ii}^{-1} G_{ii})$$

Using diagonal matrix **K,** with $K_{ii} = G_{ii}^{0.5}$

$$(\mathbf{KG(KK^{-1})V^{-1}(K^{-1}K)GK})_{ii} = 1/x_i$$
$$(\mathbf{HW^{-1}H})_{ii} = 1/x_i,$$

where $\mathbf{H} = \mathbf{KGK}$, $\mathbf{W} = \mathbf{KVK}$.

It follows that $\mathbf{H} = \begin{bmatrix} 1 & r_g \\ r_g & 1 \end{bmatrix}$ and

$$\mathbf{W^{-1}} = \frac{1}{x_1 x_2 - w^2 r_g^2} \begin{bmatrix} x_2 & -wr_g \\ -wr_g & x_1 \end{bmatrix}.$$

Thus $(\mathbf{HW^{-1}H})_{11} = \dfrac{x_2 + r_g^2 x_1 - 2r_g^2 w}{x_1 x_2 - w^2 r_g^2} = 1/x_1$

$$\frac{x_1 x_2 + r_g^2 (x_1^2 - 2wx_1)}{x_1 x_2 + r_g^2 (-w^2)} = 1$$
$$x_1^2 - 2wx_1 = -w^2$$
$$w^2 - 2x_1 w + x_1^2 = 0 = (w - x_1)^2$$
$$w = x_1$$

To prevent a change in reliability for trait 1, the residual correlation (r_e) must therefore satisfy $w = x_1$, and is derived from $\mathbf{R} = \mathbf{V} - \mathbf{G}$ as follows:

$$r_e = \frac{(w-1)}{\sqrt{(x_1-1)(x_2-1)}} r_g = \frac{(x_1-1)}{\sqrt{(x_1-1)(x_2-1)}} r_g$$

$$r_e = \sqrt{\frac{(x_1-1)}{(x_2-1)}} r_g = \sqrt{\frac{\alpha_1/\delta_1}{\alpha_2/\delta_2}} r_g = \sqrt{\frac{\delta_2/\alpha_2}{\delta_1/\alpha_1}} r_g$$

Note: $|r_e| \le |r_g|$ when $\frac{\delta_1}{\alpha_1} \ge \frac{\delta_2}{\alpha_2}$

or $\frac{\delta_1}{\delta_1+\alpha_1} \ge \frac{\delta_2}{\delta_2+\alpha_2}$

Residual correlation is thus equal to genetic correlation if reliabilities are equal, and is of relatively lower magnitude if reliability is higher for trait 1. Maintaining the same higher reliability has the consequence of increasing the lower reliability. The equation generalizes to more than 2 traits as follows:

$$r_{e(ij)} = \frac{\sqrt{(\delta/\alpha)_i (\delta/\alpha)_j}}{(\delta/\alpha)_{\max}} r_{g(ij)}$$

Residual correlations defined in this way prevent any increase in reliability for the country with highest national reliability, for the situation where genomic data sharing was 100% at the national level, but with the consequence of increasing all of the lower reliabilities. The Interbull implementation of GMACE (Sullivan and Jakobsen, 2012; 2014b) was based on residual correlations defined as (Sullivan and VanRaden, 2010):

$$r_{e(ij)} = c_{ij} \frac{\sqrt{(\delta)_i (\delta)_j}}{(\delta)_{\max}} r_{g(ij)}$$

Thus, the current implementation of GMACE will impose a correct upper limit on reliability only if heritabilities are the same in all countries, which is rarely if ever the case. This problem is easily corrected with a modified definition for GMACE residual correlations:

$$r_{e(ij)} = c_{ij} \frac{\sqrt{(\delta/\alpha)_i (\delta/\alpha)_j}}{(\delta/\alpha)_{\max}} r_{g(ij)}$$

With these new residual correlations, the upper limit for the parameter space of GMACE reliabilities is clearly defined for a given bull, when all genomic data are shared at the national level, as the maximum national reliability. A continuing consequence is that GMACE reliabilities, and corresponding variances of GMACE EBV, will be biased upwards for all countries with national reliabilities lower than the maximum. These biases can be estimated and reduced, however, by scaling the contributions of foreign data to GMACE reliabilities and EBV, as follows:

Apply GMACE equations, based respectively on 0% (*gm***I**), 100% (*gm***J**) and the specified amount of data sharing (*gm***C**). For each scale of evaluation (*i*), estimate the maximum EDC bias as: $b_i = \delta_i^{gm\mathbf{J}} - \delta_i$. With our new residual correlations: $b_i = 0$ if $(\delta/\alpha)_i = (\delta/\alpha)_{\max}$, $b_i > 0$ otherwise. The EDC bias specific to the bull is estimated as:

$$\delta_i^b = b_i * \frac{\delta_i^{gm\mathbf{I}} - \delta_i^{gm\mathbf{C}}}{\delta_i^{gm\mathbf{I}} - \delta_i^{gm\mathbf{J}}}.$$

Thus, the EDC bias will range between 0 with no genomic data sharing (C=I) and b_i with 100% data sharing (C=J). The EDC bias is easily removed by subtraction:

$$\delta_i^{gm} = \delta_i^{gm\mathbf{C}} - \delta_i^b,$$

and a corresponding multiplicative adjustment, to the GEBV contribution from foreign data, eliminates the bias of EBV ($\hat{g}_i^{gm\mathbf{C}}$) variance:

$$\hat{g}_i^{gm} = \hat{g}_i + \frac{\delta_i^{gm} - \delta_i}{\delta_i^{gm\mathbf{C}} - \delta_i} * (\hat{g}_i^{gm\mathbf{C}} - \hat{g}_i)$$

Problems caused by variance of national genomic reliabilities were reduced in GMACE, by using an international prediction model for national reliabilities (Sullivan and Jakobsen, 2014a). With the use of a parameter space to restrict GMACE results, the reliability prediction model is no longer needed.

Countries without national GEBV

Parameter space adjustments described above are for the *nobs* scales of countries providing national GEBV. Related adjustments are also needed for the *nmiss* scales of countries not providing national GEBV. The GMACE equations are easily modified to use *nobs* adjusted GEBV, rather than the unadjusted, in the GMACE multivariate genetic regression from *nobs* to *nmiss* scales.

There is no obvious way, however, to manipulate GMACE equations for a suitable multivariate adjustment of reliabilities on the *nmiss* scales. The *nmiss* reliabilities can be adjusted a posteriori, however, based on univariate methods analogous to the methods described above for the *nobs* GEBV countries.

A minimum GMACE reliability is first determined for each scale *m=1,nmiss*, as the maximum of individual foreign-country contributions (*i=1,nobs*), which could be achieved by applying GMACE with only 1 of the foreign GEBV-countries included at a time:

$$R_m^{MIN} = \max\{(R_m \mid \delta_i) = r^2_{g(im)} * \delta_i \big/ (\delta_i + 1)\}$$

This is the assumed reliability that should come from (PA-excluded portion of) GMACE if all genomic data (100%) were shared at the national level. This approximation matches well against minimum observed reliabilities from GMACE when national data sharing is very high. The EDC equivalent for this minimum GMACE reliability is denoted as:

$$\delta_m^{MIN} = \delta_m \mid R_m^{MIN} = R_m^{MIN} / (1.0 - R_m^{MIN})$$

Respective sums before (C_b) versus after (C_a) parameter space restrictions, of increases in the *nobs* EDC contributions to scale *m*, for GMACE relative to national GEBV, are used to derive scaling factors for GMACE increases above each δ_m^{MIN}. Denoting conversions between reliability and equivalent EDC as $\delta \mid R$ and $R \mid \delta$, and EDC contributions from scale *i=1,nobs* to *m=1,nmiss* as $D_{im}^x = \delta_m \mid (R_m \mid \delta_i^x)$, the required sums and EDC adjustments for scale *m* are as follows:

$$C_a = \sum_{i=1}^{nobs} (D_{im}^{gm} - D_{im})$$

$$C_b = \sum_{i=1}^{nobs} (D_{im}^{gm\mathbf{C}} - D_{im})$$

$$\delta_m^{gm} = \delta_m^{MIN} + \frac{C_a}{C_b} * (\delta_m^{gm\mathbf{C}} - \delta_m^{MIN})$$