Assessing the bias in top GPA bulls

Sargolzaei, M.[12], Chesnais, J.[1] and Schenkel, F.[2]
[1]L'Alliance Boviteq, Saint-Hyacinthe, QC, Canada
[2]University of Guelph, Centre for Genetic Improvement of Livestock, Guelph, ON, Canada

Introduction:

AI companies market a relatively small number of top bulls and therefore selecting bulls that remain at the top when proven is extremely important. The top GPA bulls tend to have lower proofs than their original GPA. This drop indicates that top GPA bulls are overestimated, i.e. there is a bias in their evaluations. A validation study was carried out to measure the average amount of bias in GPA LPI for the top 100 bulls across birth years, or for the top 35 bulls within birth year, based on their GPA in 2008 and their proof in 2012. Bias in the top bulls reflects both the average bias for all bulls and the slope in the regression of proofs over GPA. It is a useful measurement since in practice only the very top GPA bulls are marketed as young bulls and/or progeny tested.

The study was done using GPA LPI because with the advent of genomic selection, with thousands of young bulls being genotyped, GPA LPI is a primary criterion for selection of bulls in Canada. Knowledge of the bias in GPA LPI of young bulls compared to proven bulls is particularly important when deciding whether to use young sires or sons or proven sires of sons. Although CDN did report changes from a similar validation study at the CDN Open meeting in March 2012, changes in LPI were not reported.

It has also been observed that many young bulls with top GPA evaluations are the sons of bulls with a high MACE evaluation. The proofs of these sires may be overestimated and increase the average bias in GPA LPI. Furthermore, these high MACE sires are usually not used at random in Canada, so that even their domestic proof, when it becomes available, may be overestimated. The study therefore includes a look at biases in Shottle's sons, and how such biases might have been reduced, with the objective of developing approaches for other bulls that may be in the same situation in future (e.g. sons of Planet).

The average biases reported here are likely to be smaller than biases today, because all bull sires in this study were proven when their GPA was calculated. It has been shown that when parents are unproven, the GPA of progeny are less accurate compared to the GPA of progeny from proven parents (Habier et al., 2007, Genetics 177: 2389-2397). Since the proportion of top GPA bulls from unproven sires is increasing rapidly, new approaches will have to be developed to try to measure the amount of bias for such bulls when they rank towards the top.

Method and materials:

The data set was provided by CDN and consisted of 79,798 genotyped animals, 42503 SNPs. Animals that were unproven but had PA in April 2008 and were domestically proven in February 2012 were used as validation animals.

For genomic evaluation, SNP with minor allele frequency of less than 0.01 were discarded, a polygene weight of 20% was applied and the CDN blending method using traditional PA was used.

The average LPI bias was calculated for three groups of bulls: 1) the top 100 bulls based on their 2012 proofs, 2) the top 100 bulls based on their 2008 GPA for all birth years between 2003 and 2007, and 3) the top 35 bulls based on their 2008 GPA, for each birth year between 2003 and 2007. The bias for the first bull group is expected to be smaller or even positive, since in order to end up in the top 100 based on their proof, bulls are likely to have increased or dropped less than average. This situation was considered, however, for comparison with a recent article by David Selner which looked at US proofs for this type of bulls, and therefore is likely to have underestimated the real bias. The bias for groups 2) and 3) are the most interesting since they reflect a situation whereby the top bulls are selected based on their GPA but later receive a proof that may differ from it, which is what occurs in reality. Changes for group 3) bulls are for top bulls within each year, while changes for bulls in group 2) are for top selected bulls across all selection years.

The average difference between the proofs of 2,159 bulls that were proven in both 2008 and 2012 was 339 points of LPI, and this difference was used to base-adjust 2008 proofs. The biases in GPA LPI that are reported here therefore reflect changes over and above those of proven bulls. The sum of base changes between April 2008 and February 2012 was 446 points of LPI, i.e. fairly close to the average change in the proof of proven bulls.

SNP estimates and the GPA derived from them were calculated either from de-regressed proofs, as in official evaluations, or from the proof themselves. This was done because the de-regressed proofs received from CDN appear to be incorrect for some bulls. Results are normally expected to be better with de-regressed proofs, but only if the de-regression procedure functions correctly.

The effect of Shottle on the bias in GPA was also investigated. In 2008, Shottle's proof was based on MACE, and was quite high at close to 3,000 points of LPI. Bias was estimated 1) for Shottle sons; 2) when Shottle sons were removed from the validation set; 3) when Shottle's proof was removed from the estimation of SNP effects but sons were left in the validation set.

Finally, some of the changes observed in US genomic evaluations were reported for comparison with changes in the Canadian system.

Conclusions:

- The de-regression method used by CDN for official evaluations seems to produce incorrect results for some bulls. The problem can be easily seen in Table 3, for some influential bulls with high reliability and a large number of sons. Some proven bulls have de-regressed proofs for milk yield that exceed their proofs for the same trait by several thousands of kg. The problem was communicated to CDN in April 2012, at the time of this report. While the effect on the average GPA of all bulls is limited (50 points of LPI for 1,312 bulls), it is significant for the top bulls, since over-estimation increased by 230 points of LPI for the top 100 bulls when de-regressed proofs are used for the estimation of SNP effects instead of proofs. Potentially, this problem could also affect to a large extent the individual GPA of some bulls, particularly bulls whose sires have been greatly over-estimated or under-estimated.

- The effect of the de-regression problem on the GPA of top bulls, and on some individual bulls, could potentially be larger today than in 2008 given that de-regressed values appear worse in 2012 compared to 2008.

- A correct de-regression method would be expected to produce more accurate SNP estimates and predictions than the use of proofs, therefore the reduction in bias could even be larger if the current method is corrected.

- As far as the effect of Shottle and his high MACE proof in 2008 is concerned, we found that based on 2008 official evaluations, there were 58 Shottle sons in the top 100 bulls for GPA LPI. If proofs are used instead of de-regressed proofs, this number goes down to 50 and the bias is reduced from 622 to 392 points of LPI for all 100 bulls. Removing these 50 Shottle sons from the top 100 bulls for GPA LPI further reduces the bias from 392 to 159 points of LPI, showing that the bias in bulls that are not Shottle sons is much lower than for Shottle sons themselves.

- However, removing Shottle sons from official evaluations because their sire is likely to be over-evaluated would not have been an acceptable approach in 2008. It would be more desirable to find a way to reduce the bias in these sons without removing them. Therefore, an evaluation was conducted with Shottle's proof excluded from the SNP estimation group, so that SNP effects are estimated from sources other than Shottle's own proof, which was likely to be too high compared to reality. When this is done, the 90 Shottle sons receive an average GPA that is much more in line with their ultimate proofs since the average bias for these sons is reduced from 610 to 258 points of LPI. This is much better than the bias of 855 points in 2008 official evaluations for these 90 sons. As a result, the average bias for the top 100 bulls including Shottle sons is reduced from 392 to 206 points of LPI, not very far from the bias of 159 points when Shottle sons are excluded. Therefore, a potential approach for the future is to determine which specific bulls are likely to be greatly over-estimated based on their MACE evaluation or based on a non-random domestic proof (for example by looking at the number of ET of genotyped daughters), then to run two evaluations, one including the bull proof in the SNP estimation set and the other excluding it. If the GPA of the sons are very different in each case, then one could decrease the REL of the sire or even set it to zero (i.e. exclude its proof from the estimation set). The genomic evaluations of the sons would then reflect SNP effects that are based on the proofs of bulls other than their sire. This approach could be investigated for a bull like PLANET, for example. Correlations between the 2008 GPA LPI and 2012 proofs for the top young 100 bulls in 2008 were very close to zero, indicating the inability of GPA to rank the top bulls correctly among themselves. Correlations among the top 35 bulls within each year were positive but still very low, except for bulls born in 2007 but there were only 37 bulls that year.

- A rapid review of US genomic evaluations based on information made publicly available by D. Selner (Tables 10 to 13) shows that US genomic evaluations of top bulls are also seriously biased upwards for either NM$ or TPI. Considering that the true genetic SD for LPI is close to 700 points, and that it is 235 points for TPI (Tom Lawlor, pers. communication), and 199 points for $NM (AIPL site), changes of 182 points of TPI for the top 157 bulls from January 2010 to April 2012 are roughly equivalent to changes of 182*700/235 = 542 points of LPI, and for the top 20 bulls during the same period changes of 268 points of TPI are equivalent to 268*700/235 = 798 points of LPI. Approaches to reduce biases for top genomic bulls in the US system may be quite different from those in the

Canadian system. In the Canadian system, as described earlier, correcting the de-regression problem and reducing the influence of sires whose proofs are likely to be over-evaluated could potentially decrease biases by a large amount. In the US system, over-evaluation of top bulls appears more related to low b-values for some traits (for protein for example), which could be corrected through adjustments in methodology (ex. polygenic variance, one-step model, etc...).

- Regardless of the cause, the biases for the top 100 bulls are quite large (close to 1 true genetic SD) in both the Canadian and US systems. The situation is likely the same in other Interbull countries, except maybe in New Zealand where an adjustment to the genomic evaluations of top bulls was forcefully implemented following an independent review. These biases make it difficult to equitably compare proven and young bulls as potential sires of sons, since most young bulls used as sires of sons are from the top 100. In addition, the most popular young bulls marketed as sires of cows are also from the top 100, which makes it difficult for producers to decide whether they should invest in semen from proven or young bulls. Currently, the evaluation deck is heavily stacked in favour of top young genomic bulls.

- Current Interbull validation procedures miss a very simple point: the young genomic bulls that are commercialized in the industry and are now routinely used as sires of sons are not average GPA bulls, but rather those with the very top GPA. The validation parameters used by Interbull, and by member countries, i.e. square correlations and slopes for bulls in the validation set, are just as much influenced by the lowest ranking validation bulls as by the highest ranking ones, even if in practice no one cares about or uses the lowest ranking bulls. Furthermore, slopes are based on an assumption of linearity of relationship between GPA and DYD which may not be correct in practice. For example, top young bulls might be over-evaluated much more than bottom young bulls are under-evaluated. These test parameters do not reflect possible biases in the evaluations of the bulls that are actually used by the industry. Therefore, it would make sense to add one simple validation criterion to those which Interbull already has: the average bias (difference GPA vs DYD) for the top 100 bulls in the validation set, or the bias for a set % of top bulls in the validation set if their number is relatively small. This way one would actually validate the GPA of the bulls of interest, rather than those of bulls that have no effect on future genetic change in the population of member countries. Adding this simple validation criterion would eventually lead to evaluations that provide a fairer comparison between young and proven bulls, which would benefit Interbull countries and producers that use them. This could be done for more than protein yield, for example for one in each of the group of related traits submitted by member countries (production, conformation, fertility, longevity, etc...).

Results:

Table 1: Average bias when GPAs are calculated using de-regressed proofs, as per official evaluations (LPI)

	No.	Avg. 2008 GPA	Avg. 2012 proof	r^2	Bias
All bulls	1312	924	624	0.36	300
Bulls born 2003	151	559	263	0.28	296
Bulls born 2004	343	709	410	0.28	299
Bulls born 2005	368	899	624	0.35	275
Bulls born 2006	411	1215	884	0.18	331
Bulls born 2007	39	1407	1178	0.50	229

Table 2: Bias for selected animals when GPAs are calculated using de-regressed proofs, as per official evaluations (LPI)

	No.	Avg. 2008 GPA	Avg. 2012 proof	r^2	Bias
Top 100 in 2012	100	1533	1800	0.08	-267
Top 100 in 2008	100	1957	1335	0.00	622
Top 35 born 2003	35	1034	629	0.13	405
Top 35 born 2004	35	1409	928	0.02	481
Top 35 born 2005	35	1885	1267	0.00	618
Top 35 born 2006	35	2118	1242	0.02	876
Top 35 born 2007	35	1474	1250	0.43	224

Table 3: List of 10 proven bulls with largest difference between proof and de-regressed proof for MILK

Name	No. daughters	Type of proof	proof	Rel	De-reg	Rel	Diff
AEROSTAR	29100	Domestic	574	99	23964.55	67	23390.55
GOLDWYN	30571	Domestic	514	99	7112.12	57	6598.12
LEADMAN	2180	Domestic	-88	99	5107.15	89	5195.15
BARLO	3346	MACE	1053	93	5175.1	70	4122.1
MICA	10167	MACE	604	94	4189.06	65	3585.06
EMORY	50865	MACE	698	96	3855.13	38	3157.13
MASCOT	1468	Domestic	117	99	3014.07	89	2897.07
MANDEL	18694	MACE	894	98	3159.62	88	2265.62
HERALD	9620	MACE	1216	94	3459.22	76	2243.22
EMERY	9436	MACE	1187	92	2755.83	73	1568.83

Table 4: Average bias, official proofs vs GPA based on proofs (LPI)

	No.	Official (de-regressed proofs) Avg. 2008 GPA	r^2	Bias	GPA calculated from proofs Avg. 2008 GPA	r^2	Bias
All bulls	1312	924	0.36	300	877	0.41	253
Bulls born 2003	151	559	0.28	296	548	0.28	286
Bulls born 2004	343	709	0.28	299	690	0.30	280
Bulls born 2005	368	899	0.35	275	861	0.40	237
Bulls born 2006	411	1215	0.18	331	1128	0.25	245
Bulls born 2007	39	1407	0.50	229	1306	0.49	128

Table 5: Bias for selected animals, official proofs vs GPA based on proofs (LPI)

	No.	Official (de-regressed proofs) Avg. 2008 GPA	r^2	Bias	GPA calculated from proofs Avg. 2008 GPA	r^2	Bias
Top 100 in 2012	100	1533	0.08	-267	1420	0.14	-381
Top 100 in 2008	100	1957	0.00	622	1753	0.01	392
Top 35 born 2003	35	1034	0.13	405	968	0.12	351
Top 35 born 2004	35	1409	0.02	481	1315	0.10	436
Top 35 born 2005	35	1885	0.00	618	1680	0.00	367
Top 35 born 2006	35	2118	0.02	876	1884	0.03	569
Top 35 born 2007	35	1474	0.43	224	1361	0.46	119

Table 6: Average bias with and without Shottle sons included, versus bias with current official evaluations (LPI)

	No.*	All bulls Official (A)	All bulls Proofs (B)	Excl. Shottle Sons (C)	Excl. Shottle Proof (D)
All bulls	1312(90)	300	253	227	229
Bulls born 2003	151(0)	296	286	286	286
Bulls born 2004	343(0)	299	280	280	280
Bulls born 2005	368(29)	275	237	214	208
Bulls born 2006	411(61)	331	245	172	194
Bulls born 2007	39(0)	229	128	128	129
Shottle's sons	90(90)	855	610	-	258

* number of Shottle sons in parenthesis
A: All bulls - official, Shottle proof included in estimation (GPA calculated using de-regressed proofs)
B: All bulls, Shottle proof included in estimation (GPA calculated using proofs)
C: Shottle proof included in estimation but his sons removed from validation (GPA calculated using proofs)
D: All bulls, Shottle proof removed from estimation (GPA calculated using proofs)

Table 7: Bias for selected animals with and without Shottle sons included, versus bias with current official evaluations (LPI)

	No.	All bulls Official (A)	All bulls Proofs (B)	Excl. Shottle Sons (C)	Excl. Shottle Proof (D)
Top 100 in 2012	100	-267	-381	-423	-417
Top 100 in 2008	100	622	392	159	206
Top 35 born 2003	35	405	351	351	364
Top 35 born 2004	35	481	436	436	452
Top 35 born 2005	35	618	367	94	128
Top 35 born 2006	35	876	569	254	286
Top 35 born 2007	35	224	119	119	121

A: All bulls - official, Shottle proof included in estimation (GPA calculated using de-regressed proofs)
B: All bulls, Shottle proof included in estimation (GPA calculated using proofs)
C: Shottle proof included in estimation but his sons removed from validation (GPA calculated using proofs)
D: All bulls, Shottle proof removed from estimation (GPA calculated using proofs)

Table 8: Bias for Shottle's son within selected animals (LPI)

	All bulls Official (A)*	All bulls Proofs (B)*	Excl. Shottle Proof (D)*
Top 100 in 2012	392(11)	118(11)	-223(11)
Top 100 in 2008	914(58)	642(50)	397(18)
Top 35 born 2005	861(23)	554(20)	413(5)
Top 35 born 2006	1043(27)	733(20)	353(9)

* the number of Shottle sons is in parenthesis
A: All bulls - official, Shottle proof included in estimation (GPA calculated using de-regressed proofs)
B: All bulls, Shottle proof included in estimation (GPA calculated using proofs)
D: All bulls, Shottle proof removed from estimation (GPA calculated using proofs)

Table 9: USDA-Holstein prediction accuracy

Trait[a]	Bias[b]	b	REL (%)	REL gain (%)
Milk (kg)	-64.3	0.92	67.1	28.6
Fat (kg)	-2.7	0.91	69.8	31.3
Protein (kg)	0.7	0.85	61.5	23.0
Fat (%)	0.0	1.00	86.5	48.0
Protein (%)	0.0	0.90	79.0	40.4
PL (months)	-1.8	0.98	53.0	21.8
SCS	0.0	0.88	61.2	27.0
DPR (%)	0.0	0.92	51.2	21.7
Sire CE	0.8	0.73	31.0	10.4
Daughter CE	-1.1	0.81	38.4	19.9
Sire SB	1.5	0.92	21.8	3.7
Daughter SB	- 0.2	0.83	30.3	13.2

[a]PL = productive life, CE = calving ease and SB = stillbirth.
[b]2011 deregressed value – 2007 genomic evaluation.

Table 10: Selner's data set

GPTA to proof	**Number of bulls***	**Average change in TPI**	**Average change in in $NM**
Jan 2010 to Dec 2011	121	-170	-138
Jan 2010 to April 2012	157	-182	-149

*** Bulls marketed on GPTA in Jan 2010 that now have a proof**

Table 11: Selner's data set - Jan 2010 to April 2012

Choice of bulls	Trait	Average in 2010	Average in 2012	Difference
Top 10 bulls in Jan 2010	TPI	2,311	2,080	-231
	Net Merit	718	480	-238
Top 10 bulls in Apr 2012	TPI	2,259	2,178	-81
	Net Merit	663	622	-41
Top 20 bulls Jan 2010 to Dec 2011	TPI	2,228	1959	-269
	Net Merit	668	503	-165
Top 20 bulls Jan 2010 to Apr 2012	TPI	2,273	2,005	-268
	Net Merit	691	489	-202

Table 12: Selner's data set - Jan 2010 to April 2012

	Young bulls (out of 157)	Once proven
Number of bulls above 2,100 TPI	46	10
Number of bulls above 600 NM$	34	5

Table 13: Selner's comparison

Top bulls	Fat eval. (lbs)		Protein eval. (lbs)	
	Jan 2010 (GPTA)	Dec 2011 (Proof)	Jan 2010 (GPTA)	Dec 2011 (Proof)
Top 10 young genomic bulls in Jan 2010*	83	56	54	37
Top 10 proven bulls in Jan 2010	92	86	68	66

* selection of these bulls was likely done for more than just fat or protein