Effect of Number of Bulls on Sire Variance Estimation

Filippo Miglior[12], Peter G. Sullivan[3] and Brian J. Van Doormaal[2]
[1]Agriculture and Agri-Food Canada, Dairy and Swine Research and Development Centre, Lennoxville, QC, Canada
[2]Canadian Dairy Network, Guelph, ON, Canada
[3]CGIL, University of Guelph, Guelph, ON, Canada

Introduction

In Interbull evaluations sire variance estimates are used to account for different scales of proof expression in each country. Sire variances are estimated by REML using all bulls that have daughters in at least 10 herds, by country. This group includes locally sampled bulls and imported bulls first sampled abroad. Also each bull' ancestors, whose proofs are not based on daughter performance, are included in the sire variance estimation process. Previous research in Canada (Van Doormaal *et al.*, 1999; Van Doormaal & Miglior, 2000) and Italy (Cassandro *et al.*, 1997) has shown that various countries have heterogeneous trends in sire variance estimates over time. In addition, each country's relative sire variance estimate is known to have an important impact on international sire rankings computed by Interbull using MACE.

Through field and simulation studies, Miglior *et al.* (2001) concluded that the current estimation method used by Interbull overestimates sire variances due to the inclusion of 2nd country proofs of selected bulls and do not properly account for heterogeneity of sire variances over time and across countries. Biases in sire variance estimates could be removed, and variance heterogeneity effectively accounted for, by estimating sire variances for each country using only bulls born in the most recent 5 years.

The investigation of Miglior *et al.* (2001) considered the Holstein populations of ten countries that had the highest numbers of proven bulls over time. The effects of smaller population size on sire variance estimation were not considered. The objective of this study was to investigate the effect of number of bulls on sire variance estimation and the impact of using various strategies on the stability of sire variance estimation in small populations.

Identification of Small Populations

Interbull files from August 2001 were used to identify the smaller populations (breed-country combination that had less than 100 proven bulls in total). Proven bulls in each country were bulls with at least 10 herds in that country for protein yield, stature or SCS. Frequencies of bulls by year of birth were computed for each breed-trait-country combination. The total number of bulls might have been smaller than the one actually used by Interbull in official runs, since the latter also included unofficial bulls that were not distributed in the Interbull released files.

Small populations from the August 2001 production run are reported in Table 1. The Jersey breed had the highest number of countries with less than 100 proven bulls. The smallest population was Italian Jersey with only 11 bulls with daughters in at least 10 herds. Two small populations were also found in the Holstein breed: Slovenia and French Red. Ayrshire and Simmental did not have populations with less than 100 bulls. Distributions of bulls by year of birth for the small populations are reported in Table 2. Of particular note was that for the Jersey breed there were many years with no bulls, for all three of the small population. For example, Italy had only 7 years of bulls with gaps in '84, '86 and '90-'91 years. Data in Tables 1 and 2 are specific to the number of bulls for the production evaluation. Analysis was also performed for type and udder health evaluations. Type evaluation was excluded from subsequent analyses because a different sire variance program has been used for Interbull evaluation of type traits and a yearly estimate could therefore not be estimated.

Table 1. Small populations, with less than 100 bulls evaluated by Interbull for production in August 2001evaluation.

Breed	Country	No. of bulls
Brown Swiss (BSW)	Canada (CAN)	55
	The Netherlands (NLD)	39
Guernsey (GUE)	Canada (CAN)	55
	South Africa (ZAF)	25
Holstein (HOL)	Slovenia (SLO)	63
	French Red (FRR)	82
Jersey (JER)	The Netherlands (NLD)	23
	Italy (ITA)	11
	Germany (DEU)	18

Table 2. Number of bulls, by year of birth, for small populations evaluated by Interbull for production in August 2001evaluation.

Birth	BSW		GUE		HOL		JER		
year	CAN	NLD	CAN	ZAF	SLO	FRR	NLD	ITA	DEU
79	1		1	1			2		
80	2		1	3			2		
81		2	4	2			4		
82	5	3	1	5			6	1	1
83	2		3	3			3	1	
84	5	4	3	1	4	3			
85	7	4	3	2	2	11	1	1	2
86	3	3	5	1	5	6			
87	7	2	4	4	5	5		2	
88	4	4	11	1	9	7	1	4	3
89	2	2	1		4	9		1	3
90	4	2	2	1	7	12	1		1
91	3	6	4	1	3	3	1		
92	2	2	2		5	6		1	2
93	2	3	5		9	6			4
94	2	2	2		5	6			1
95	2		3		5	6	2		1
96	1					2			
97	1								

Stability of Sire Variance Estimates over time

Sire variance estimation log files, kindly provided by Interbull, were used to study the stability of sire variance estimate through two Interbull routine runs in small populations: a) May 2001 (new populations included and first evaluation for udder health), and b) February 2002 (most recent official run at the time of this study). Sire variance was estimated for milk, fat and protein yield and SCS using five different groups of bulls: 1) all bulls and ancestors (**Interbull**), 2) bulls born in the last five years (**Last 5-yr**), 3) bulls born in the last six years (**Last 6-yr**), 4) bulls born in the last 10 years (**Last 10-yr**), and 5) bulls born in the last 5 or more years, as needed to include at least 20 bulls (**20 bulls**).

Results for Slovenia were excluded because of significant changes in their national evaluation methodologies between May 2001 and February 2002. Absolute changes in sire SD were computed across the two runs for each population and trait. In Table 3 number of years and bulls are reported for each small population across the two Interbull runs.

A summary of average changes of sire SD over time is reported in Table 4 for the five different groups of bulls. The magnitude of changes was greater when only recent born bulls were used (**Last 5-yr**, **Last 6-yr** and **20 bulls**). When all bulls (**Interbull**) or only bulls born in the last 10 years (**Last 10-yr**) were used, changes were generally smaller. When only recently born bulls were used, larger changes were expected due to recent trends in sire variance. Heterogeneity of changes was evident across breeds, with the largest changes in the smaller Jersey populations and smaller changes in the Brown Swiss and Holstein populations. Smaller populations clearly had larger changes when results were pooled across breeds by population size.

A separate analysis was performed on five large Holstein populations (Canada, France, Germany, The Netherlands and United States) using results from the same two Interbull runs (May 2001 and February 2002). The average absolute change in sire SD across breeds, traits and countries when all bulls and ancestors (**Interbull**) were included was equal to .7%. With only bulls born in the last 5 years (**Last 5-yr)**, the change was equal to 1.2%. Thus, the magnitude of changes in sire SD was much bigger in smaller compared to larger populations, as expected.

Table 3. Distribution of bulls and birth years by MACE run.

	Feb 2002		May 2001	
	Years	*Bulls*	*Years*	*Bulls*
BSW				
CAN	16	56	17	53
NLD	13	39	13	38
GUE				
CAN	18	62	17	56
ZAF	12	26	12	25
HOL				
FRR	13	83	13	81
JER				
NLD	10	23	8	20
ITA	8	14	8	13
DEU	9	18	8	17

Table 4. Average absolute change of sire SD across Interbull runs for small populations averaged across traits.

	Interbull	Last 5-yr	Last 6-yr	Last 10-yr	20 bulls
Within breed					
BSW	1.4%	4.2%	4.7%	1.7%	3.8%
GUE	2.4%	10.6%	12.9%	5.2%	11.9%
HOL	1.2%	4.4%	9.2%	2.5%	4.4%
JER	6.7%	10.0%	12.9%	6.2%	7.2%
By size					
< 20 bulls	6.7%	10.0%	12.9%	6.2%	7.2%
20-60 bulls	1.8%	7.2%	8.5%	3.3%	7.6%
> 60 bulls	1.2%	4.4%	9.2%	2.5%	4.4%
Overall	**3.4%**	**7.8%**	**10.1%**	**4.2%**	**7.0%**

Minimum Number of Bulls per Year

Canadian Holstein bull proofs from the November 2000 Interbull run were used to determine a minimum number of bulls per year necessary to have an accurate estimate of sire variance. In the last 5 years 1827 bulls were born. Sire SD using this group of bulls was 8.12 kg of protein. The complete data set included around 400 bulls per year and 130 born in the last year (less than 30% of the complete group by year). Independent random samples of actual bulls were drawn to estimate sire variance with decreasing numbers of bulls per year (15, 10, 8, 5, 4, 3 and 2) and a smaller number in the last year. The number of random samples was limited by the size of the sample since each bull was in one sample only (independent samples). Table 5 summarizes the estimates of sire SD from the random samples. Sire SD estimates did not change when at least 20 bulls were used. When less than 20 bulls were used sire SD tended to be smaller, although not significantly different from 8.12 kg (p>.05).

Magnitude of standard deviation of sire SD estimates increased exponentially as the number of bulls decreased (Figure 1). A much larger SD was expected with less than 20 bulls. With number of bulls greater than 20 the curve tended to plateau.

Table 5. Estimation of sire SD by decreasing sample size.

Last 5 years	N	No. of samples	Sire SD estimate
All bulls born in the last 5 years	1827		8.12
23 bulls per year and 7 in the last year	99	17	8.09 ± .57
18 bulls per year and 6 in the last year	78	21	8.12 ± .35
15 bulls per year and 5 in the last year	65	26	8.12 ± .72
12 bulls per year and 4 in the last year	52	33	8.10 ± .81
10 bulls per year and 3 in the last year	43	37	8.09 ± .83
8 bulls per year and 2 in the last year	34	49	8.09 ± .97
5 bulls per year and 1 in the last year	21	77	8.10 ± 1.25
4 bulls per year and 1 in the last year	17	100	8.05 ± 1.44
3 bulls per year and 1 in the last year	13	131	7.94 ± 1.60
2 bulls per year and 1 in the last year	9	199	7.88 ± 2.05

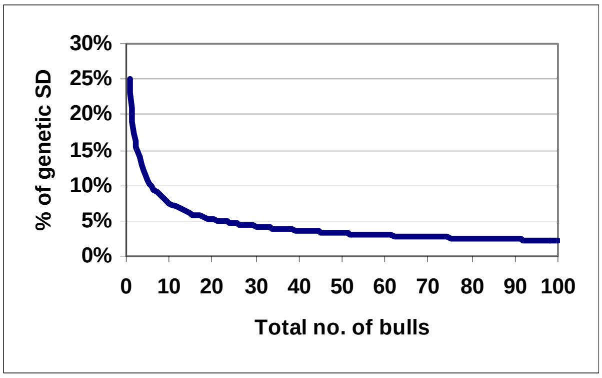

Figure 1. Variation of sire SD estimates by number of bulls included (protein kg).

Conclusions

Even in a large breed like the Holstein there were countries with a relatively small number of bulls per year of birth. For example, Slovenia and French Red had a total of 63 and 82 bulls, respectively, over a 13-year period. The smallest populations were in the Jersey breed: Italy (11 bulls), Germany (18 bulls) and The Netherlands (23 bulls). Estimation of sire variances in small populations was susceptible to changes across subsequent runs because of the small number of bulls included in the estimation. Magnitude of changes was inversely proportional to population size with largest changes occurring when less than 20 bulls were included. Using only the last 5 years yielded higher changes than when all bulls were included. This was expected since the **Last 5-yr** estimate accounts more effectively for trends in sire variance occurring over time as newly proven bulls get added. For very small populations (less than 20 animals) the **Last 5-yr** strategy can produce SD estimates with very high PEV. Thus, it might be more appropriate to modify the recommended **Last 5-yr** strategy to apply the **20 bulls** strategy for populations with less than 20 bulls in the last 5 years.

References

Cassandro, M., Miglior, F., Carnier, P., Bittante, G., Canavesi, F., Santus, E. & Banos, G. 1997. Effect of standardization of within country-year sire variance of deregressed proofs on international evaluations. *Interbull Bulletin 16*, 16.

Miglior, F., Sullivan, P.G. & Van Doormaal, B.J. 2001. Alternative strategies for estimation of country sire variance in international evalution of dairy bulls. *Interbull Bulletin 27,* 84.

Van Doormaal, B.J., Kistemaker, G.J. & Sullivan, P. G. 1999 Heterogeneous variances of Canadian bull EBV over time. *Interbull Bulletin 22*, 141.

Van Doormaal, B.J. & Miglior, F. 2000. Trends in sire variance estimates by birth year. *Interbull Bulletin 25,* 70.