

Preliminary Analysis of Mendelian Sampling Terms for Genetic Evaluation Validation

Filippo Miglior^{1,2}, Peter G. Sullivan³ and Brian J. Van Doormaal²

¹Agriculture and Agri-Food Canada, Dairy and Swine Research and Development Centre, Lennoxville, QC, Canada

²Canadian Dairy Network, Guelph, ON, Canada

³CGIL, University of Guelph, Guelph, ON, Canada

Introduction

Currently Interbull implements two validation tests on submitted data: a genetic trend test and a sire standard deviation test. The genetic trend test (Bonaiti *et al.*, 1994) comprises three methods. In the first one, proofs derived from a multiple lactation repeatability model are compared to proofs derived from first lactation data only, under the assumption of unity correlations across lactations in the repeatability model. Estimated genetic trends with both models are expected to yield similar results. In the second method, the within-sire daughter yield deviations (DYD) are analyzed by daughter birth year and are expected to remain stable over time. The third method analyzes variations of official proofs over time by regression.

The methods used for genetic trend test are becoming less appropriate for most situations. The first method is not suitable for current genetic evaluation methodologies that do not use a multiple lactation repeatability model. DYD are seldom available for the second method because they are no longer required for the MACE evaluations and are difficult to estimate for a TDM. The third method does not consider changes to genetic evaluation methodologies, which are continuously being updated. The sire standard deviation test does not assess the validity of a national evaluation but focuses on the degree of change in sire SD for consecutive MACE runs. The efficacy of both the genetic trend

and the sire SD tests raises the need for new validation tests for genetic evaluations.

Recently, trends in average and variance of Mendelian Sampling (MS) terms have been used to validate national evaluations (Van Doormaal *et al.*, 1999; Canavesi & Samoré, 2000; Gengler *et al.*, 2000; Miglior & Van Doormaal, 2000). The theoretical expectation is that trends in mean and variance of MS should remain constant over time. The overall objective of this project is to develop validation tests based on trends of MS over time, which can be applied nationally by each genetic evaluation centre on both cows and bulls, and internationally by Interbull for bulls.

National Validation Tests

The first step of the project was to analyze trends of MS for bulls and cows from the Canadian Test Day Model. National production proofs of Holsteins from August 2001 were used in the analysis. MS were computed for cows born after 1990 that had known parents and performance records, and for bulls born after 1984 that had daughters in at least 10 herds. MS for milk, fat and protein yield were computed from EBV for 1st, 2nd, 3rd and combined lactations. A total of 4,666 bulls and 880,859 cows were analyzed (Table 1). Proofs from August 1999 and 2000 were also analyzed to investigate the effect of evaluation year on MS trends.

Table 1. Distribution of bulls and cows by year of birth from August 2001 run.

Birth year	Bulls	Cows
84	226	
85	261	
86	297	
87	378	
88	386	
89	393	
90	436	
91	462	61,828
92	445	84,554
93	461	103,185
94	400	118,808
95	393	129,635
96	128	126,100
97		131,049
98		125,700
Total	4,666	880,859

Trends in average MS, by trait and lactation respectively, are reported for cows in Figures 1 and 2, and for bulls in Figures 3 and 4. The overall trends were generally nonlinear with increasing trends in recent years. There were no major differences across traits (Figure 1 and 3) or lactations (shown for protein only in Figures 2 and 4). Trends were similar in pattern for cows relative to bulls, but were much larger in magnitude for bulls.

Trends in SD of MS for cows were very similar across traits: flat until '96 and then a decrease for the last 2 years (Figure 5). First lactation SD had a flat trend over time (Figure 6), while 2nd lactation SD had a decrease in the last year, and 3rd lactation SD in the last 2 years, due to the amount of information for each animal. Animals born in '98 had only 1st lactation data, those born in '97 1st and 2nd lactation, and only animals born in '96 or earlier had information for all lactations (Figure 7). Thus, the trend in SD of MS for the combined proof was very similar to the trend in average cow reliability, as shown in Figure 8.

Trends in SD of MS were different for bulls than for cows. For bulls there was an increasing trend until '91 and then a steady decrease until the last year of birth (Figure 9). Trends seemed to be affected by second crop information, which was available for

bulls born before '92. Bulls born in '91 had the largest proportions of second crop daughters with only early first lactation records. The pattern was quite evident when consecutive August runs were analyzed for 1999, 2000 and 2001: the peak shifted ahead each year as the highest proportions of daughters with early first lactation records moved to the next bull birth year (Figure 10). While the decreasing pattern in recent years seemed to follow the decreasing trend in reliability, it was not the case for the peak for second crop bulls. Further research is needed to analyze other possible causes.

Validation Tests for Interbull

Interbull has access to bull proofs and pedigree for a sire model that could be used for validation of national evaluations provided by each country. Trends in MS for the sire model of MACE were studied using Canadian data submitted to Interbull for the November 2000 run. The MACE program for sire variance estimation was modified to output estimates of MS for all bulls.

In the MACE sire model, MS is computed as:

$$MS = EBV_{\text{bull}} - [\frac{1}{2}EBV_{\text{sire}} + \frac{1}{4}EBV_{\text{MGS}} + \frac{1}{4}EBV_{\text{MGD}}]$$

All EBV_{MGD} are phantom group solutions. The time edit, which excluded EBV from animals born before 1983 for the Interbull November 2000 run, increases the impact of phantom group solutions on MS. For example, 50% of Canadian bulls did not have a sire with an EBV and 82% of bulls did not have a MGS with an EBV. Thus, many ancestor EBV, used to compute MS, were not based on daughter performance.

In Figure 11 average MS of bulls were plotted over time for protein yield (November 2000 evaluation). The trend for the MACE model were close to zero with a noticeable increase in the last year. The trend for the animal model was mostly flat but slightly negative and tended to increase only for most recently born bulls. Trends in SD of MS had a very similar pattern, for both the animal and MACE models, with a peak in 1990 and sharp decreases for the rest of the period (Figure 12). Adjustment for PEV of MS

(Sullivan, 1999) had only a small effect on the SD trend.

Discussion and Conclusion

Each national genetic evaluation centre should be capable of checking MS trends for female and male populations. Interbull could only check trends in the bull population and would require validation tests for the cow population in the same way as it is currently done with the genetic trend validation. Validation tests based on MS trends could be applied to any trait analyzed regardless of the model. The PEV of MS may affect MS trends, and should be accounted for in the validation tests. The PEV are available from the Interbull program to estimate sire variances (Sullivan, 1999) and can therefore be accounted for in MS trend validation tests used by Interbull. The PEV of MS would not be available for the animal model, however, due to computational requirements and approximations may therefore be required.

Assuming that PEV of MS are known or can be approximated with reasonable accuracy, trends in mean and SD of MS could be used to validate national genetic evaluations of both cows and bulls. Regression analysis could be used to test for trends that deviate significantly from zero. Subsets of populations could be compared. Animals with positive MS could be compared to animals with negative MS by year of birth. Average squared MS should be equal for positive relative to negative MS animals. Local bulls could be analyzed separately from imported bulls, as SD trends for imported sires will be strongly affected by selection.

If a suitable approximation for PEV of MS is developed, a combination of the above options could be applied nationally for MS of cows. The sire estimation program in MACE could be modified to produce a series of tests based on adjusted MS of bulls, as outlined above. The same program could be used at the national level before submitting data to Interbull. A modified program could provide validation test results automatically, and could be used by Interbull to accept or refuse incoming national genetic evaluations.

References

- Bonaiti, B., Boichard, D., Barbat, A. & Mattalia, S. 1994. Three methods to validate the estimation of genetic trend in dairy cattle. *Interbull Bulletin* 10.
- Canavesi, F. & Samoré, A.B. 2000. Guidelines for national genetic evaluations and tools for data quality checks. Interbull Workshop, Verden, Germany. *Mimeo*.
- Gengler, N., Dusseldorf, T., Wiggans, G.R., Wright, J.R. & Druet, T. 2000. Joint estimation of variances and effects in the US Jersey type evaluation system. Interbull Workshop, Verden, Germany. *Interbull Bulletin* 26, 34.
- Miglior, F. & Van Doormaal, B.J. 2000. Impact of second country proofs in genetic evaluation. *Interbull Bulletin* 25, 57-59.
- Sullivan, P.G. 1999. REML estimation of heterogeneous sire (co)variances for MACE. *Interbull Bulletin* 22, 146.
- Van Doormaal, B.J., Kistemaker, G.J. & Sullivan, P.G. 1999 Heterogeneous variances of Canadian bull EBV over time. *Interbull Bulletin* 22, 141.

Figure 1. Average cow Mendelian Sampling by trait (Combined EBV – Aug 2001)

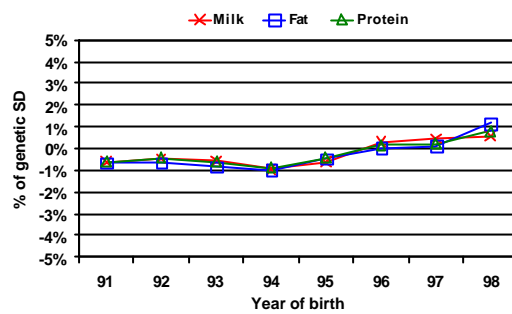


Figure 2. Average cow Mendelian Sampling by lactation (Protein – Aug 2001)

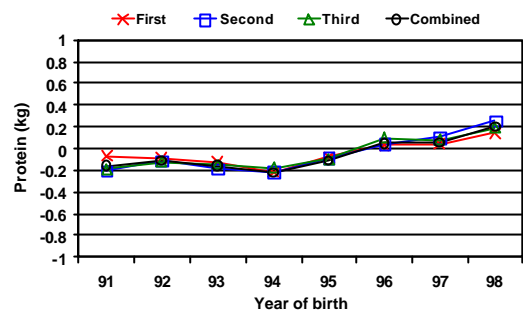


Figure 3. Average bull Mendelian Sampling by trait (Combined EBV – Aug 2001)

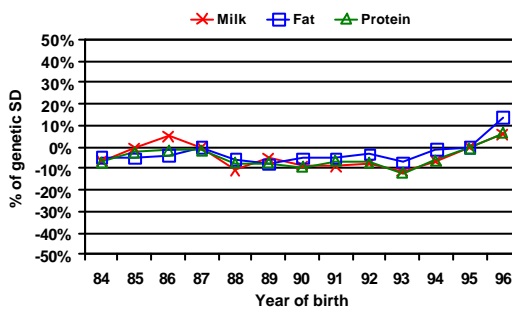


Figure 4. Average bull Mendelian Sampling by lactation (Protein – Aug 2001)

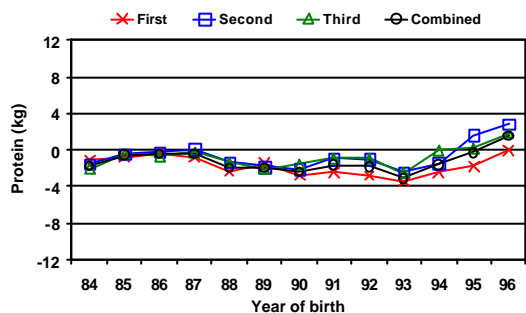


Figure 5. SD of cow Mendelian Sampling by trait (Combined EBV – Aug 2001)

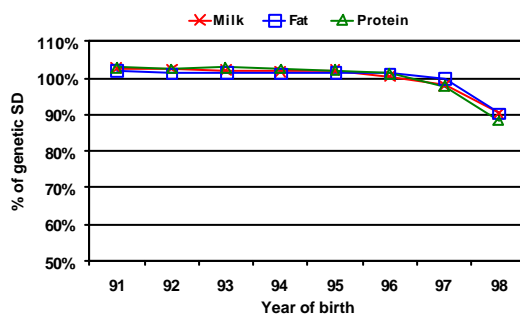


Figure 6. SD of cow Mendelian Sampling by lactation (Protein – Aug 2001)

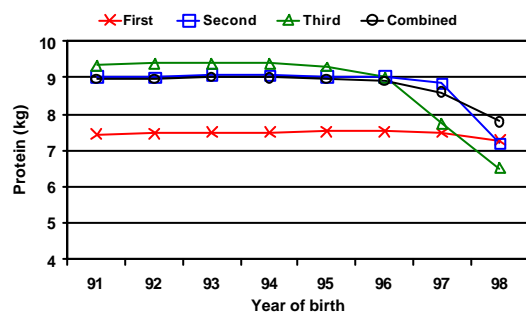


Figure 7. Number of average test days per cow by parity (Aug 2001)

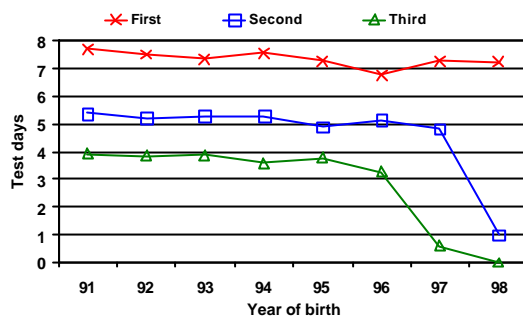


Figure 8. SD of cow Mendelian Sampling and reliability (Protein – Aug 2001)

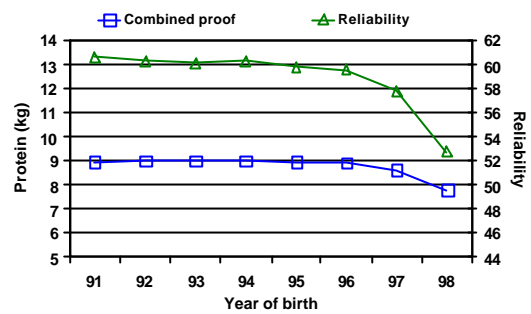


Figure 9. SD of bull Mendelian Sampling by trait (Combined EBV– Aug 2001)

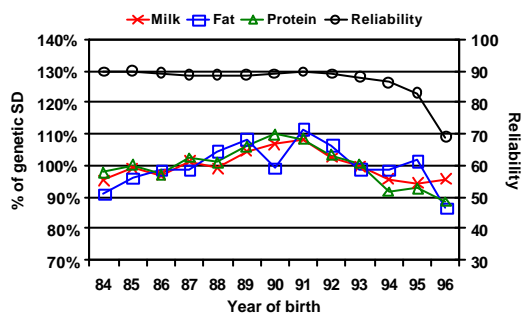


Figure 10. Quadratic trends, from regression, in SD of bull MS (August consecutive runs)

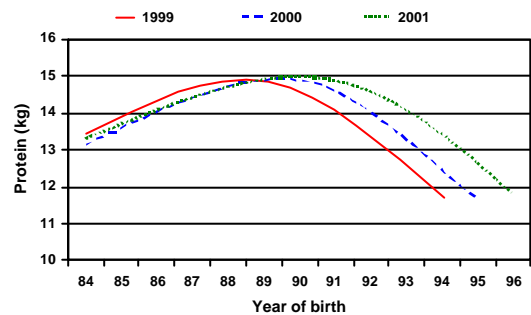


Figure 11. Average bull Mendelian Sampling (Combined EBV - Nov 2000)

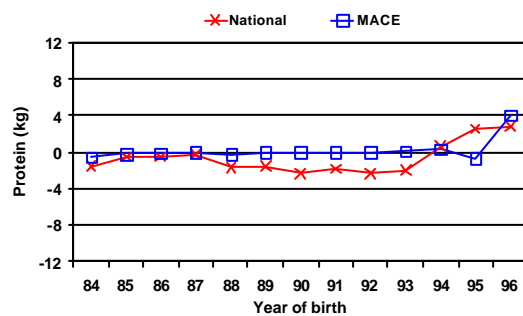


Figure 12. SD of bull Mendelian Sampling (Combined EBV - Protein - Nov 2000)

