Use of a Weighted Random Regression Test-Day Model to Better Relate Observed Somatic Cell Score to Mastitis Infection Likelihood

N. Gengler [1,2] and P. Mayeres [2]
[1]*National Fund for Scientific Research, B-1000 Brussels, Belgium*
[2]*Animal Science Unit, Gembloux Agricultural University, B-5030 Gembloux, Belgium*

Abstract

Current methods to analyze somatic cell counts or scores do not differentiate the origins of observed values, they rely on the simple hypothesis that a "higher" value is linked to an intra-mammarian infection. This hypothesis may not be totally valid, but until now proposed alternatives were not usable in large scale genetic evaluations. A simple modification of currently used random regression test-day models is presented that weight observed somatic cell scores by a weight that expresses mastitis infection likelihood. Results showed that despite high correlation with breeding values obtained without the weight, differences that were observed were highly related to mastitis results in Nordic countries. Sire breeding values obtained with this method were also submitted to the March 2003 INTERBULL test-run. They showed the largest average correlations of all populations with mastitis results from Nordic populations.

The proposed method is currently implemented and used in routine in the genetic evaluation for somatic cell score in the Walloon Region of Belgium.

Introduction

The current methodology for estimating breeding values for somatic cell count (**SCC**) or scores (**SCS**) does not account for the potential difference in origins of observed SCC. As a matter of fact the simple hypothesis that a "higher" SCC is linked to an intra-mammarian infection event is not necessarily valid. Despite evidence that selection for lower SCC is decreasing mastitis incidence without weakening natural defenses (e.g., Rupp and Boichard, 2000) research in the direction of differentiation SCC according to their potential origins could be beneficial. However few scientific work focussed on alternative methods to analyze SCC. Detilleux and Leroy (2000) proposed an interesting idea, a mixed normal mixture model, but for technical reasons they could not use it in large scale real-life situations. The objective of the present paper was to show that the general idea to take into account likely intra-mammarian infection status when analyzing SCC or SCS can be achieved in an alternative manner using a weighted random regression test-day model (**RRTDM**). A practical implementation in the Walloon RRTDM for SCS is shown.

Materials and Methods

Method

Through the modeling of SCS by a RRTDM, every SCS observation can be considered as the sum of the expected SCS value given the model and the model solutions and the deviation from this expected value. This deviation can then be considered a indicator for the presence of an intra-mammarian infection event. Elvinger *et al.* (1991) described the use of a similar deviation but they lacked the possibility to use a RRTDM. Test-day SCS can then be weighted iteratively by an indicator function reflecting mastitis infection likelihood. This method allows then to differentiate observed SCS values. Even if this approach does only show limited similarities to the mixed normal mixture model

in the sense that both methods try to separate using infectional status, both try to achieve the same goal by computing a expected probability that a given SCS is linked to an intra-mammarian infection event.

Weighted RRTDM

A typical RRTDM for SCS can be written as

$$\mathbf{y = Xb + Qa + Qp + e}$$

where **y** is the vector of daily SCS, **b** is a vector of fixed effects, **p** a vector of permanent environmental random regression, **a** a vector of additive genetic random regressions. One can define at every iteration a weight that is a direct function of the observed standardized residuals $e_s = \frac{e}{\sigma_e^2}$. The choice of this function is at the present stage quiet arbitrary. It should only reflect that a standardized residuals approaching $-\infty$ equals the lowest possible weight and approaching $+\infty$ equals the highest possible weight. A logical choice was therefore a sigmoid function and the following function chosen:

$$\gamma = \frac{1+\alpha\beta}{1+\alpha\exp(-e_s)}$$

where α was put to 1 assuming a symmetric distribution and β to 1.6506 in order to allow the mean weight γ being 1.

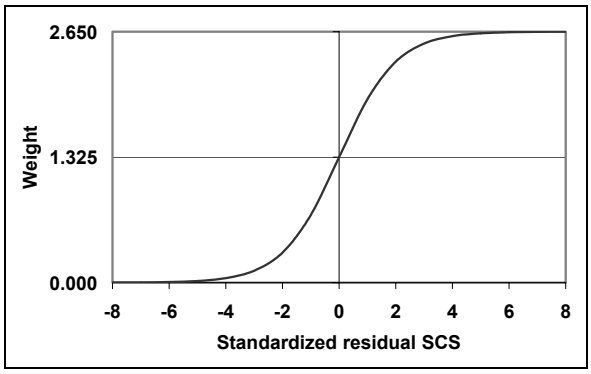

Figure 1. Evolution of the weight γ in function of standardized residual SCS.

Data

SCC data was provided by the Walloon Breeding Association (AWE) who manages performance recording data in the Walloon Region of Belgium. Data edition was done to keep records in the first three lactation occurring between 5 and 365 DIM and to exclude unlikely ages for a given lactation or gestation lengths. Additionally, recorded SCC had to be at least 10000 cells per ml and be below 10 million cells. SCS were computed using the formula by Ali and Shook (1980) $SCS = [\log_2(SCC/100000)] + 3$ SCS below 0.1 were put to 0.1. As shown in Figure 1 this commonly used transformation allows to get a traits that has a nearly normal distribution (data from August 2003), especially in the second and third lactation. Table 1 gives additional details of the data used in the routine run of August 2003.

Table 1. Number of test-day records used in August 2003 and observed means and standard-deviations (STD) for SCS.

		SCS	
Lactation	Records	Mean	STD
1	4050039	3.00	1.60
2	3065412	3.41	1.71
3	2249564	3.73	1.76

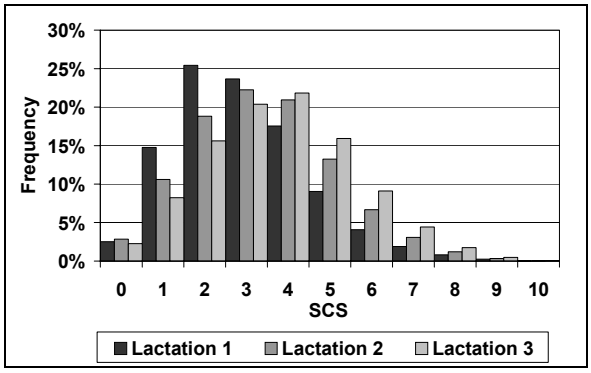

Figure 2. Distribution of first, second and third lactation TD SCS/.

Genetic evaluation model

The genetic evaluation model was a multi-lactation RRTDM similar to the one used for milk, fat and protein that was described by Auvray and Gengler (2002). There were only two major modifications, no common herd environment effects were defined as previous research showed that this effect is very small for SCS and the model was only multilactation, not multitrait.

(Co)variance components

(Co)variance components used were derived in a similar manner as described by Auvray and Gengler (2002). Table 2 shows that genetic values for SCS were similar for the three lactation even if phenotypic values were rather poorly correlated.

Table 2. Heritabilities for SCS on the diagonal, genetic correlations above, phenotypic correlations below.

	Lactation		
Lactation	1	2	3
1	0.10	0.95	0.88
2	0.31	0.13	0.95
3	0.34	0.42	0.16

Computation and expression of breeding values

Solving of the weighted RRTDM was obviously more difficult than for a regular RRTDM. Experience showed that keeping first weights equal to 1 until a high level of convergence allow a smooth solving. Estimated breeding values (EBV) were computed from random regressions expressed as a mean value over 305 lactation days and the three lactation. Values were also expressed as deviations from a common genetic base that was put to 3 for the mean of all cows born in 1995 with records. The value of 3 was chosen as it represents the first lactation mean. This presentation was preferred to relative values as breeders in the Walloon Region are used to SCS as these values are also used as management tools. In the following study EBV computed using the weights are called EBVw, EBV obtained without weighting EBVo.

Comparison of results with and without iterative weights

In order to allow comparison of results using the August 2003 run data a weighted and a not weighted run were computed and EBV obtained. Additionally genetic evaluations from INTERBULL routine runs for Holstein and Red Holstein populations in Denmark, Sweden and Finland were obtained to test if changes in our results were likely linked to mastitis.

INTERBULL test-run March 2003

Results obtained in March 2003 using the weighted RRTDM described here were submitted to the INTERBULL test-run. Under the rationale that the proposed method does a better job than ordinary RRM or lactation models the correlation obtained by INTERBULL should be at the higher edge of the range of correlation among populations using SCS or SCC and the direct mastitis evaluations of the Nordic Holstein or Red Holstein populations : Denmark, Sweden, and Finland.

Results and Discussion

Comparison of results with and without iterative weights

As expected correlation among EBVw and EBVo was very high with values of 0.986 for the 706 bulls send to INTERBULL. Table 3 gives details about the distribution of the SCS breeding values for these bulls and the difference observed.

Table 3. Details of the EBV for the 706 bulls send to INTERBULL for the August 2003 routine run (EBVw = with weights, EBVo = without weights.

	Mean	Std	Min	Max
EBVo	3.01	0.36	1.98	4.34
EBVw	3.03	0.40	2.04	4.51
EBVo – EBVw	-0.03	0.07	-0.33	0.16

The most interesting detail was that despite the high correlation there were substantial differences in the observed values.

Table 4. Correlation of the difference EBVo-EBVw with mastitis results from Nordic populations for different minimum levels of EDC in our data.

Minimum		Population			
EDC	N	DNK	DNR	SWE	FIN
All	636	0.48	0.46	0.33	0.26
50	550	0.49	0.48	0.33	0.25
150	330	0.54	0.52	0.38	0.24

DNR = Danish Red Holstein

Table 4 shows the correlation of the difference between EBVo and EBVw and genetic evaluations for mastitis from INTERBULL routine runs for Holstein and Red Holstein populations in Denmark, Sweden and Finland. These results are expressed on a scale where higher values indicate less mastitis incidence. Therefore a positive correlation with EBVo-EBVw means that this difference is positively related to mastitis resistance. The higher correlation with the Danish populations could be due to the fact that their models are multitrait somatic cells and mastitis. Also our method tries to detect clinical and sub-clinical mastitis, where usual direct mastitis recordings are (mostly?) clinical.

INTERBULL test-run March 2003

Correlations obtained by INTERBULL in the March 2003 test-run are summarized in Table 5. The results showed that our average correlation was the highest with the Nordic mastitis results compared to all other populations using SCS or SCC. Even if one should not overemphasis the value of a correlation, as observed value can be due to a great number of factors, this seems to indicate that the method worked reasonably well.

Table 5. Comparison of average correlation obtained for mastitis evaluations in Nordic populations with SCC or SCS results in other populations.

	All other populations			
WAL	Mean	Std	Min	Max
0.65	0.59	0.03	0.51	0.64

Conclusions

Current genetic evaluation systems for SCC or SCS do not account for the potential different origins of observed values. This paper presents a very simple way to do this, at least approximately. The method presented here is simple and can be used easily in current RRTDM settings. It was used successfully in the Walloon genetic evaluation system providing breeding values that were better predictors of mastitis.

Implications

The method presented in this paper has the potential for further developments. First, the method used to compute iterative weights is definitely still rather crude, but given the results has some potential for further improvement. Also; if direct milk recording systems keep track of mastitis events at a given test-day this information can be directly feed into the system considering the maximum weight for this test. Finally the current way to express results may not be optimal and alternative expression of overall EBV could provide better indicator functions for mastitis.

Acknowledgments

The authors gratefully acknowledged the support of the Walloon Breeding Association (**AWE**) and the Walloon Regional Ministry of Agriculture (Projects : RW1009, D31-1039).

Nicolas Gengler, who is Chercheur Qualifié of the National Fund for Scientific Research (Brussels, Belgium), acknowledges support through Grant 2.4507.02 F (2) of the National Fund for Scientific Research.

References

Ali, A.K.A. & Shook, G.E. 1980. An optimum transformation for somatic cell concentration in milk. *J. Dairy Sci. 63,* 487-490.

Auvray, B. & Gengler, N. 2002. Feasibility of a Walloon test-day model and study of its potential as tool for selection and management. *INTERBULL Bulletin 29,* 123-127.

Detilleux, J. & Leroy, P.L. 2000. Application of a mixed normal mixture model for the estimation of Mastitis-related parameters. *J. Dairy Sci. 83,* 2341-2349.

Elvinger, F., Littell, R.C., Natzke, R.P. & Hansen, P.J. 1991. Analysis of somatic cell count data by a peak evaluation algorithm to determine inflammation events. *J. Dairy Sci. 74,* 3396-3406.

Rupp, R. & Boichard, D. 2000. Relationship of early first lactation somatic cell count with risk of subsequent first clinical mastitis. *Livest. Prod. Sci. 62,* 169-180.