55th Meeting of the European Association for Animal Production,
Bled, Slovenia, 5. -9. September 2004
Session on developments in biometrical methods: G6-10

Comparison of two methods to approximate reliabilities of breeding values from test day model

I. Strandén and E.A. Mäntysaari
MTT Agrifood Research Finland, 31600 Jokioinen, Finland

INTRODUCTION

Reliability or its square root, accuracy, is often reported by breeding organizations and required by Interbull along with estimated breeding values. In theory, reliability is calculated as a function of genetic variance and elements of the inverse of the coefficient matrix of the mixed model equations that are used to estimate breeding values. In practice, approximation methods are used due to heavy computational load in calculation of the inverse.

Jamrozik et al. (2000) presented an approximation method, named JSJ method hereinafter, that accounts both observation and relationship information. Liu et al. (2002), and Tier and Meyer (2004) have presented recently new methods with comparisons to true reliability and the JSJ method. According to Liu, the JSJ method severely overestimated reliabilities of bulls with first lactation test day records only. However, Tier and Meyer (2004) found R-squared values above 97.4% between the true reliability and the JSJ method. The application was for beef cattle where reliabilities for weight of breeding values estimated by a random regression model were calculated. Both Liu et al. (2002) and Tier and Meyer (2004) presented approximation methods that were better than the JSJ method, although both methods were quite different.

Aim of this paper is to compare reliabilities by the JSJ method and the approximation method by Misztal and Wiggans (1988) to true reliabilities. Dairy cattle data with test day records and random regression test day model are used.

MATERIAL AND METHODS

Let **C** be coefficient matrix of the mixed model equations (MME). Inverse of the MME coefficient matrix can be partitioned into blocks

$$\mathbf{C}^{-1} = \begin{bmatrix} \mathbf{C}_{bb} & \mathbf{C}_{bu} \\ \mathbf{C}_{ub} & \mathbf{C}_{uu} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{C}^{bb} & \mathbf{C}^{bu} \\ \mathbf{C}^{ub} & \mathbf{C}^{uu} \end{bmatrix}$$

where superscript b refers to non-genetic effects and u refers to the genetic effects. Let breeding value for animal i be $a_i = \mathbf{k}'\mathbf{u}_i$ where vector $\mathbf{u}_i$ has breeding values of animal i and vector **k** represents the weights. Reliability of breeding value a_i is calculated as

$$r_i^2 = 1 - \frac{\mathbf{k}'\mathbf{C}_i^{uu}\mathbf{k}}{\mathbf{k}'\mathbf{G}\mathbf{k}}$$

where **G** is the genetic (co)variance matrix, and $\mathbf{C}_i^{uu}$ is diagonal block for animal i in $\mathbf{C}^{uu}$.

Methods to calculate reliabilities

Reliability was calculated by two approximation methods and by inverting the coefficient matrix using sparse matrix techniques (the inverse method). The approximation methods were the JSJ methods, and a modified Misztal and Wiggans (1988) approximation, named hereinafter MW method. Both approximation methods had 2 steps. The first step was the same, and there the coefficients of the non-genetic effects in MME were accounted by absorbing their contributions to the $\mathbf{C}_{uu}$ matrix. Let R_i^2 be reliability calculated from the first step. In the second step relationship information was accounted. After the first step, both approximate methods used only a single value per animal in the second step.

Originally the MW method was presented to calculate reliabilities for all individuals in an animal model. However, the method can be used for multiple trait models by considering an aggregate single trait breeding value. The important step is to calculate reliability for an animal without genetic information as was done for single trait analysis presented in their paper. Reliability of an animal with observations is transformed to a single (aggregate) trait without genetic information or 'effective number of records':

$$d_i = \alpha \frac{R_i^2}{1 - R_i^2}, \tag{1}$$

where R_i^2 is reliability after accounting for non-genetic effects, $\alpha = (1 - h^2)/h^2$, $h^2 = \mathbf{k'Gk}/\sigma_p^2$ and σ_p^2 is the phenotypic variance of the aggregate trait. This formula for d_i follows by assuming (Misztal and Wiggans, 1988)

$$R_i^2 = 1 - \frac{\alpha}{\alpha + d_i} \tag{2}$$

In the second step, relationship information assuming animal is accounted by iteratively traversing pedigree as presented in their paper.

The JSJ method has a similar formula to (1). The JSJ method uses equivalent number of progeny (ENP) which is the number of daughters animal should have to get the same reliability level as achieved from the animal's phenotypic records. ENP for animal *i* is

$$\mathrm{ENP}_i = \lambda \frac{R_i^2}{1 - R_i^2}$$

where $\lambda = (4 - h^2)/h^2$. This ENP is then used to incorporate information from relationships by traversing pedigree twice as presented in Jamrozik et al. (2000).

Data set

Data set with first lactation test day records of milk, fat, and protein were sampled from the Finnish Ayrshire data base. The sample had records from 20 herds. Breeding values were assumed to be estimated by a random regression test day model used in Finland (Lidauer et al., 2000) but only having first lactation records, i.e., variance components of random effects for the first lactation were equal to those in the Finnish test day model. The model was simplified to have only within herd effects: herd test day, permanent environment, and animal genetic. Permanent environment and animal genetic effects were modeled by a random regression curve with 4 coefficients: 2nd order polynomial and a Wilmink term. Rank reduction had been applied to decrease sizes of the animal permanent environment and genetic (co)variance matrices from 12 to 6.

The data had 3478 animals. Cows were from 20 herds, and number of cows with records was 2260. There were 12698 milk records, 6152 protein records, and 6154 fat records. Reliabilities were calculated for 305 day breeding values by taking 10 days with 30 days interval from day 15, i.e., weights for estimated breeding values were $\mathbf{k} = \sum_{\substack{i=15 \text{ to } 285 \\ \text{step } 30}} \phi_w(i)$ where $\phi_w(i)$ has the regression function coefficients of the trait (milk, protein or fat) for additive genetic effect calculated at day *i*.

RESULTS AND DISCUSSION

Descriptive statistics between reliabilities calculated by the inverse and the approximation methods show good agreement (Table 1). Almost all statistics indicate that the MW method is slightly better than the JSJ method. However, the differences are minimal and both methods agree well with the true reliabilities calculated by the inverse method.

Table 1. Correlations, mean differences ($\text{Mean}_{\text{diff}}$), maximum differences ($\text{Max}_{\text{diff}}$), and mean squared error (MSE) for reliabilities by the approximation methods and true values from inverse of the coefficient matrix.

	Correlation (%)		$\text{Mean}_{\text{diff}}$ ($\times 10^{-3}$)		Max_{diff}		MSE ($\times 10^{-5}$)	
Trait	MW	JSJ	MW	JSJ	MW	JSJ	MW	JSJ
Milk	99.98	99.94	2.51	-4.91	-0.055	-0.077	2.17	8.31
Protein	99.96	99.92	-6.01	0.70	-0.049	-0.066	6.92	5.81
Fat	99.96	99.91	2.66	-2.18	-0.051	-0.072	2.49	4.55

Differences in reliabilities by the approximation methods and true reliability were plotted, and the traits that had most deviations from true reliability are in Figures 1 and 2. Figure 1 shows the tendency for the MW method to overestimate large reliabilities. Still, most of the reliabilities were overestimated only by value of at most 0.02. The JSJ method, on the other hand, seemed to underestimate low reliabilities (Figure 2). Overestimation of the high reliabilities was less than by the MW method. It seemed that bull reliabilities were often overestimated by the MW method, i.e., there was positive bias for bull reliabilities. For the JSJ method, most of the underestimated reliabilities were from bulls. The MW method gave always higher or equal reliabilities than the JSJ method. Our results for the JSJ method agree with those by Tier and Meyer (2004) but we did not find the substantial positive bias in reliabilities of bulls reported by Liu et al. (2000).

CONCLUSIONS

Approximation methods for reliability that account relationships are computationally feasible for large data sets and seem to give reasonably good estimates of the true reliability.

REFERENCES

Jamrozik, J., Schaeffer, L.R., Jansen, G.B. 2000. Approximate accuracies of prediction from random regression models. Livest. Prod. Sci. 66:85-92.

Lidauer, M., Mäntysaari, E. A., Strandén, I, Pösö, J. 2000. Multiple trait random regression test-day model for all lactations. Interbull bulletin 25:81-86.

Liu, Z., Reinhardt, F., Reents, R. 2002. The multiple trait effective daughter contribution method applied to approximate reliabilities of estimated breeding values of a random regression test day model for genetic evaluation in dairy cattle. In: Proc. 7th World Congr. Genet. Appl. Livest. Prod., Aug. 19-23, 2002, Montpellier, France. CD-ROM communication no. 20-15.

Misztal, I., Wiggans, G.R. 1988. Approximation of prediction error variance in large-scale animal models. J. Dairy Sci. :27-32.

Tier, B., Meyer, K. 2004. Approximating prediction error covariances among additive genetic effects within animals in multiple-trait and random regression models. J. Anim. Breed. Genet. 121:77-89.

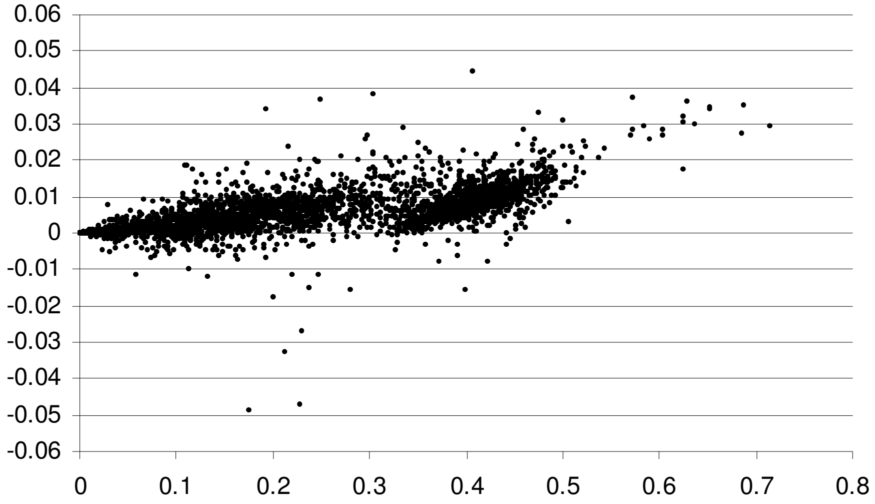

Figure 1. Differences in reliabilities for protein by the Misztal and Wiggans (1988) method and true values based on inverse of the mixed model coefficient matrix. True value in the X axis.

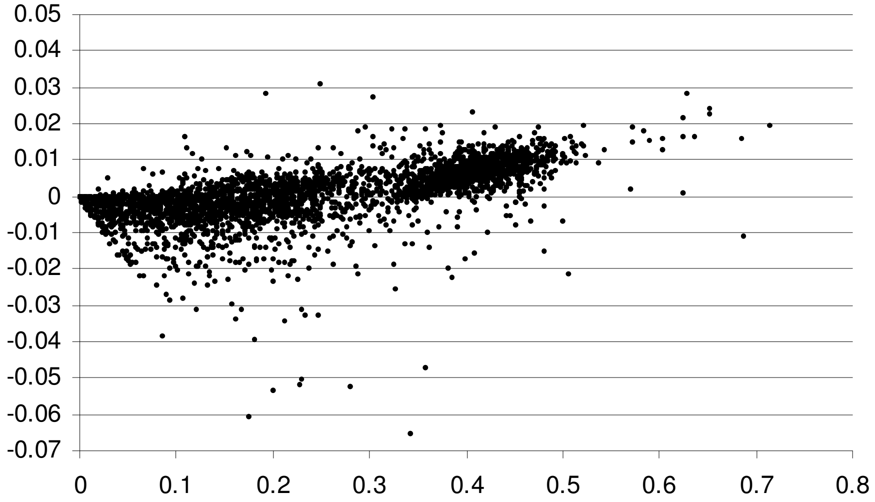

Figure 2. Differences in reliabilities for protein by the Jamrozik et al. (2000) method and true values based on inverse of the mixed model coefficient matrix. True value in the X axis.