THE CALCULATION OF COW AND DAUGHTER YIELD DEVIATIONS AND PARTITIONING OF GENETIC EVALUATIONS WHEN USING A RANDOM REGRESSION MODEL

R. A. Mrode and G. J. T. Swanson

Animal Data Centre, Greenways Business Park, Chippenham, Wiltshire, SN15 1BN, England

INTRODUCTION

The derivation and calculation of daughter yield deviation (DYD) using an animal model was reported by VanRaden and Wiggans (1991). DYD, which is a measure of unregressed daughter performance, has become very important in dairy cattle research. It was initially the variable of choice for international evaluations by Interbull but due to the inability of several countries to calculate DYD, de-regressed proofs were used (Sigurdsson and Banos, 1995). In a recent study, Madsen et al (2001) demonstrated that when connections among countries are few, the use of de-regressed proofs for low to moderately heritable traits resulted in a downward bias in genetic parameters estimated using AI-REML. However, the use of DYDs gave estimates similar to the true values. Moreover, DYDs are commonly used in dairy cattle studies aimed at detecting quantitative trait loci based on the grand-daughter design (Weller, 2001). The recent trend in dairy cattle genetic evaluations is towards application of random regression models (RRM) using test day (TD) records. The calculation of DYDs using a RRM have not been reported. VanRaden and Wiggans (1991) also gave simplified equations that explained animal evaluations in terms of contributions from various sources of information. Similar equations have not been presented for a RRM to date.

This paper outlines the calculation of DYDs and cow yield deviations when using a RRM. Equations similar to those derived by VanRaden and Wiggans (1991) are presented for RRM. Calculations are illustrated with evaluations for somatic cell count (SCC) using a RRM.

MATERIALS AND METHODS

The RRM for the evaluation of $\log_e$ SCC in the first three lactations as different traits is :

$$\mathbf{y_i} = \mathbf{X_i b_i} + \mathbf{Q_i a_i} + \mathbf{Z_i pe_i} + \mathbf{e_i} \qquad (1)$$

Where $\mathbf{y}_i$ is the vector of TD records for the i^{th} parity; $\mathbf{b}_i$ the vector of fixed effects consisting of herd-test-days, fixed lactation curves nested within age by season subclasses within parity ; $\mathbf{a}_i$ and $\mathbf{pe}_i$ are vectors of random regressions for animal and permanent environmental effects respectively and $\mathbf{e}_i$ the vector of the residual effect. The matrices $\mathbf{X}_i$, $\mathbf{Q}_i$ and $\mathbf{Z}_i$ are the incidence and covariable matrices. $\mathbf{Q}_i$ and $\mathbf{Z}_i$ are matrices of orthogonal polynomials of order 2 and 3 respectively for days in milk (DIM). It was assumed that the variances of $\mathbf{a}$, $\mathbf{pe}$ and $\mathbf{e}$ were $\mathbf{G} \otimes \mathbf{A}^{-1}$, $\mathbf{P} \otimes \mathbf{I}$ and $\mathbf{R} \otimes \mathbf{I}$ respectively ; where $\mathbf{G}$ and $\mathbf{P}$ are genetic and permanent environmental covariance matrices respectively, among regression coefficients and traits ; $\mathbf{A}$ = additive numerator relationship among animals and $\mathbf{R}$ = diagonal matrix for residual variances of size 4 per parity. The evaluation for the Holstein breed using equation (1) involved 1287930 cows with 11920882, 6342101 and 3787796 TD records for SCC in parities 1, 2 and 3 respectively.

Calculating cow yield deviations. Test day records contribute to random regressions for animal effect through the yield deviations (**YD**) and can be calculated as

$$\mathbf{YD} = (\mathbf{Q'R^{-1}Q})^{-1}\mathbf{Q'R^{-1}}(\mathbf{y} - \mathbf{X\hat{b}} - \mathbf{Zp\hat{e}}) \qquad \textbf{(2)}$$

YD is a vector of weighted regressions of the animal's TD yields adjusted for all effects other than additive genetic effect, on orthogonal polynomials for DIM. The YD vector can be used to generate actual yield deviations for any DIM and in explaining cow evaluations.

Partitioning random regression coefficients for animals. Equations for the random regression coefficients for animals are

$$(\mathbf{Q'R^{-1}Q} + \mathbf{A^{-1}G})\mathbf{\hat{a}} = \mathbf{Q'R^{-1}}(\mathbf{y} - \mathbf{X\hat{b}} - \mathbf{Zp\hat{e}}) \qquad \textbf{(3)}$$

Let $\mathbf{\hat{u}} = ½\ \mathbf{\hat{a}}$, random regression coefficients for animal effect expressed on Predicted Transmitting Ability (PTA) scale. Transferring the left-diagonal terms of $\mathbf{A}^{-1}$ in equation (3) to the right side of the equation (VanRaden and Wiggans, 1991) gives

$$(\mathbf{Q'R^{-1}Q} + \mathbf{G^{-1}}\alpha_{\mathbf{anim}})\mathbf{\hat{u}_{anim}} = \mathbf{G^{-1}}\alpha_{\mathbf{par}}(\mathbf{\hat{u}_{sire}} + \mathbf{\hat{u}_{dam}}) +$$

$$(\mathbf{Q'R^{-1}Q})\mathbf{YD}/\mathbf{2} + \mathbf{G^{-1}}\sum\alpha_{\mathbf{prog}}(\mathbf{\hat{u}_{prog}} - \mathbf{0.5\hat{u}_{mate}})$$

Where α_{par} = 1, 2/3 or 1/2 if both, one or neither parents are known respectively and α_{prog} = 1 if animal's mate is known and 2/3 if unknown. Note that $\alpha_{\mathbf{anim}} = 2\,\alpha_{\mathbf{par}} + 0.5\,\alpha_{\mathbf{prog}}$.

$$(\mathbf{Q'R^{-1}Q} + \mathbf{G^{-1}}\alpha_{\mathbf{anim}})\mathbf{\hat{u}_{anim}} = 2\mathbf{G^{-1}}\alpha_{\mathbf{par}}\mathbf{PA} + (\mathbf{Q'R^{-1}Q})\mathbf{YD}/2 + .5\mathbf{G^{-1}}\sum\alpha_{\mathbf{prog}}(2\mathbf{\hat{u}_{prog}} - \mathbf{\hat{u}_{mate}}) \qquad \textbf{(4)}$$

where **PA** = parent average. Dividing both sides of the equation by $\mathbf{Q'R^{-1}Q} + \mathbf{G^{-1}}\alpha_{\mathbf{anim}}$ gives

$$\mathbf{\hat{u}_{anim}} = \mathbf{W_1PA} + \mathbf{W_2}(\mathbf{YD}/2) + \mathbf{W_3PC} \qquad \textbf{(5)}$$

with $\mathbf{PC} = \sum\alpha_{\mathbf{prog}}(2\mathbf{\hat{u}_{prog}} - \mathbf{\hat{u}_{mate}})$. Note that $\mathbf{W_1} + \mathbf{W_2} + \mathbf{W_3} = \mathbf{I}$ and are of the order of the orthogonal polynomials for animal effect. These matrices of weights give the relative weight given to contributions from the random regressions for **PA**, **YD** and progeny. Equation (5) is useful in explaining the evaluations for animals in terms of contributions from different sources of information and how these contributions vary with DIM. The above weights apply to random regression coefficients. However, PTA for 305-day SCC or milk yield (PTA_{305}) may be required. Then equation (5) becomes

$$\mathbf{PTA_{305anim}} = \mathbf{V_1T'PA} + \mathbf{V_2T'}(\mathbf{YD}/2) + \mathbf{V_3T'PC} \qquad \textbf{(6)}$$

where **T** is a vector of order n (n = order of polynomial for animal effects) with elements t_i calculated as

$$t_i = \sum_{i=1}^{n}\sum_{j=1}^{305} q_{ij}$$

where q_{ij} are elements of **Q** and $\mathbf{V_i} = \mathbf{D_iW_i}$. Matrix **D** is diagonal with $d_i = t_i / \sum_{i=1}^{n} |\,t_i\,|$

Calculating daughter yield deviations. PC in (5) is a regressed measure of progeny performance. DYD is a more independent and unregressed measure of daughter performance. For the daughter of a bull, with no progeny of her own, equation (5) becomes

$$\hat{\mathbf{u}}_{\mathbf{prog}} = \mathbf{W}_{\mathbf{1}_{\mathbf{prog}}}\mathbf{PA} + \mathbf{W}_{\mathbf{2}_{\mathbf{prog}}}(\mathbf{YD}/2) \qquad \textbf{(7)}$$

Let **PC** be expressed as in equation (4) :

$$\mathbf{PC} = 0.5\mathbf{G}^{-1}\sum\alpha_{\mathbf{prog}}(2\hat{\mathbf{u}}_{\mathbf{prog}} - \hat{\mathbf{u}}_{\mathbf{mate}}) \qquad \textbf{(8)}$$

Substituting equation (7) into equation (8) gives :

$$\mathbf{PC} = 0.5\mathbf{G}^{-1}\sum\alpha_{\mathbf{prog}}(\mathbf{W}_{\mathbf{1}_{\mathbf{prog}}}\hat{\mathbf{u}}_{\mathbf{anim}} + \mathbf{W}_{\mathbf{1}_{\mathbf{prog}}}\hat{\mathbf{u}}_{\mathbf{mate}} + \mathbf{W}_{\mathbf{2}_{\mathbf{prog}}}\mathbf{YD} - \hat{\mathbf{u}}_{\mathbf{mate}})$$

Since these daughters have no offspring, $\mathbf{W}_{\mathbf{1}_{\mathbf{prog}}} = \mathbf{I} - \mathbf{W}_{\mathbf{2}_{\mathbf{prog}}}$, therefore

$$\mathbf{PC} = 0.5\mathbf{G}^{-1}\sum\alpha_{\mathbf{prog}}((\mathbf{I} - \mathbf{W}_{\mathbf{2}_{\mathbf{prog}}})\hat{\mathbf{u}}_{\mathbf{anim}} + \mathbf{W}_{\mathbf{2}_{\mathbf{prog}}}(\mathbf{YD} - \hat{\mathbf{u}}_{\mathbf{mate}})) \qquad \textbf{(9)}$$

Substituting equation (9) into equation (4) and moving all terms involving $\hat{\mathbf{u}}_{\mathbf{anim}}$ to the left hand side gives :

$$(\mathbf{Q'R}^{-1}\mathbf{Q} + 2\mathbf{G}^{-1}\alpha_{\mathbf{par}} + 0.5\mathbf{G}^{-1}\sum\mathbf{W}_{\mathbf{2}_{\mathbf{prog}}}\alpha_{\mathbf{prog}})\hat{\mathbf{u}}_{\mathbf{anim}} = 2\mathbf{G}^{-1}\alpha_{\mathbf{par}}\mathbf{PA} + (\mathbf{Q'R}^{-1}\mathbf{Q})\mathbf{YD}/2 +$$

$$0.5\mathbf{G}^{-1}\sum\mathbf{W}_{\mathbf{2}_{\mathbf{prog}}}\alpha_{\mathbf{prog}}(\mathbf{YD} - \hat{\mathbf{u}}_{\mathbf{mate}})$$

Dividing both sides of the equation by the coefficient of $\hat{\mathbf{u}}_{\mathbf{anim}}$, gives

$$\hat{\mathbf{u}}_{\mathbf{anim}} = \mathbf{M}_1(\mathbf{PA}) + \mathbf{M}_2(\mathbf{YD}/2) + \mathbf{M}_3(\mathbf{DYD}) \qquad \textbf{(10)}$$

where $\mathbf{DYD} = \sum\mathbf{W}_{\mathbf{2}_{\mathbf{prog}}}\alpha_{\mathbf{prog}}(\mathbf{YD} - \hat{\mathbf{u}}_{\mathbf{mate}})/\sum\mathbf{W}_{\mathbf{2}_{\mathbf{prog}}}\alpha_{\mathbf{prog}}$ and $\mathbf{M}_1 + \mathbf{M}_2 + \mathbf{M}_3 = \mathbf{I}$

Equation (10) is the same as derived by VanRaden and Wiggans (1991) for the univariate situation. However, in the TDM situation, the **DYD** in (10) are regression coefficients and actual DYD can be generated for any DIM.

RESULTS AND DISCUSSION

Daily YD at different DIM are calculated and plotted (Figure 1) for 2 cows, A and B, with PTAs of –13% and 15% respectively for SCC in the first parity using equation (2).

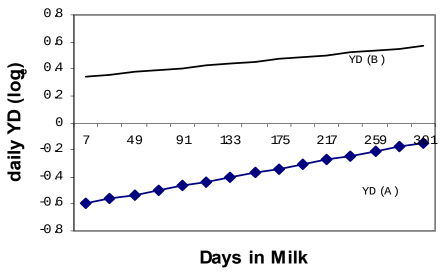

Figure 1. Yield deviations (YD) for SCC at different stages of lactations for two cows A and B

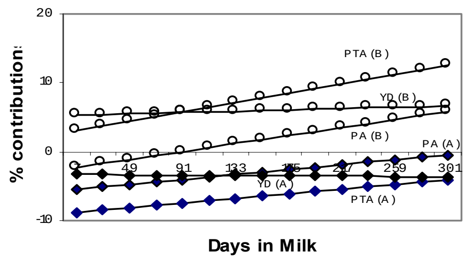

Figure 2. Contributions of parents (PA) and yield (YD) to SCC PTA at different stages for cows A and B

The trend in daily YD corresponds closely with their PTA for SCC. Thus Cow A, with a negative PTA (-13%), has negative YDs throughout the lactation in contrast with cow B, with positive YDs and PTA (+15%). he daily YDs averaged over the whole parity were –0.37 and 0.45 respectively for cows A and B. When expressed on the same scale as the PTAs, these YDs should be multiplied by 100. These YDs could be useful in explaining cow evaluations.

The daily PTAs, calculated using equation (6), for cows A and B and contributions from PA and YD, are in Figure 2. For Cow A, daily contributions of YDs were rather stable throughout the whole lactation but PA varied from –6% at the beginning of the parity to –0.5% at the end of the parity. For cow B, YD contribution was higher for most part of the parity. The contributions of daily PA averaged over the whole parity to 305-day SCC PTA of cows A and B were –6% and 3% respectively. Corresponding contributions from YD were –7% and 12%.

The correlations between DYDs calculated using equation (10) and PTAs for SCC for bulls from the RRM and current animal model (AM) based on completed lactations are in Table 1. As the number of daughters increases, the correlation between DYDs and PTAs increases, almost becoming unity as number of daughters reaches 200.

Correlations were similar for both RRM and AM with large numbers of daughters. However, with less than 50 daughters, the correlations were higher in the RRM and this may be due to a better correction of environmental factors in the RRM.

Table 1. Correlations between DYDs and PTAs in both RRM and Animal Models (AM)

No of bulls	No of daughters	Random Regression Model		AM
		First parity only	All 3 parities	
10556	≤ 20	0.74	0.73	0.44
2786	21-50	0.99	0.97	0.92
270	100-150	0.99	0.98	0.96
700	> 200	0.99	0.99	0.99

CONCLUSION

Similar to AM evaluations based on completed lactations, random regression effects for animals from RRM can be partitioned in terms of contributions from parents (PA), records (YD) and progeny (PC). The simple equations are useful in explaining evaluations from RRM for any DIM. The calculation of DYD is straightforward, as in the situation with animal model evaluation; and it can be generated for any stage of lactation.

REFERENCES

Sigurdsson, A. and Banos, G. (1995) *Acta. Agric. Scand.* **45** : 209-217
VanRaden, P. M. and Wiggans, G. R. (1991) *J. Dairy Sci.* **74** : 2737-2746
Weller, J. I. (2001) Quantative trait loci analysis in animals. *CAB International, UK*