Derivation of Weights on Daughter and Pedigree Information for Estimating Breeding Values of Progeny Tested Bulls Using a Test Day Model

Z. Liu, F. Reinhardt, R. Reents and L. Dopp

United Datasystems for Animal Production (VIT), Heideweg 1, D-27283 Verden, Germany

Introduction

Three sources of information contribute to the estimation of breeding value (BV) of an animal: pedigree information, offspring records and own performance. In case of milk production traits of dairy cattle, only daughter and pedigree information contribute to BV estimation for progeny tested bulls. Reliability of estimated breeding values (EBV) provides a measure of the total amount of information from both sources. For animal breeders, in addition to reliability, it is also important to know relative weights on the two sources of information associated with EBV. The objectives of this study are to derive weights on pedigree and daughter information for estimating BV of progeny tested bulls using a test day model (Reents et al. 1995, 1998), and to compare the weight estimates under different scenarios.

Statistical Methods

Statistical model for genetic evaluation: A test day model has been used in routine genetic evaluations of dairy production traits for German Holstein population since August 1998 (Reents et al. 1995, 1998). In this approach, first three lactations are modelled as genetically different traits and test day records of the same lactation are treated as repeated

records. For a cow with records, the test day model is

$$y = X\beta + Za + Zp + e$$

where

- *y* is a vector of test day records of her first three lactations,
- β represents fixed effects affecting test day records, including herd-test-date-parity and the shape of lactation curve effects, X is the design matrix for β ,
- *a* is a vector of random additive genetic effects of first three lactations,
- *p* is a vector of random permanent environmental effects of first three lactations,
- \mathbf{Z} is the design matrix for \boldsymbol{a} or \boldsymbol{p} , and
- *e* is a vector of random residual effects.

Let G_o , P and R_0 represent (co)variance matrices for additive genetic, permanent environmental and residual effects of the first three lactations, and G_o^{-1} , P^{-1} and R_0^{-1} their respective inverse matrices. For daughter *i* of a bull, the diagonal block (D_i) and right hand sides (RHS) pertinent to this daughter in mixed model equations (MME) are:

$$\boldsymbol{D}_{i} = \begin{bmatrix} \boldsymbol{Z}' \boldsymbol{R}^{-1} \boldsymbol{Z} + \boldsymbol{a}^{ii} \boldsymbol{G}_{o}^{-1} & \boldsymbol{Z}' \boldsymbol{R}^{-1} \boldsymbol{Z} \\ \boldsymbol{Z}' \boldsymbol{R}^{-1} \boldsymbol{Z} & \boldsymbol{Z}' \boldsymbol{R}^{-1} \boldsymbol{Z} + \boldsymbol{P}^{-1} \end{bmatrix}, \text{ and } \begin{bmatrix} \boldsymbol{Z}' \boldsymbol{R}^{-1} \boldsymbol{y} \\ \boldsymbol{Z}' \boldsymbol{R}^{-1} \boldsymbol{y} \end{bmatrix}$$

where $\mathbf{R}^{-1} = \mathbf{R}_o^{-1} \otimes \mathbf{I}$, and a^{ii} is the diagonal element corresponding to daughter *i* in the inverse of a relationship matrix. The inverse of the diagonal block above will be

used in later derivation as: $C_{i} = D_{i}^{-1} = \begin{bmatrix} C_{i}^{11} & C_{i}^{12} \\ C_{i}^{21} & C_{i}^{22} \end{bmatrix}$ *Absorptions:* For deriving weights on daughter and pedigree information for estimating BV of test bulls, daughters' additive genetic and permanent environmental effects are absorbed in the bull, resulting in a diagonal block for bull's additive genetic effects:

$$a^{bb}G_{o}^{-1} - \sum_{i=1}^{n_{d}}G_{o}^{-1}C_{i}^{11}G_{o}^{-1}$$

Corresponding RHS is:

$$\sum_{i=1}^{n_d} G_o^{-1} (C_i^{11} + C_i^{12}) Z_i' R^{-1} y_i$$

where subscript *i* refers to daughter *i*, n_d is the number of daughters with records, and a^{bb} is the diagonal element of the bull in the inverse of the relationship matrix.

If test day records of the bull dam are also available for genetic evaluation, then the permanent environmental effects of the dam are to be absorbed into her own additive genetic effects. The diagonal block and RHS of additive genetic effects of the dam after the absorption are:

$$a^{dd}G_{o}^{-1} + (I - Z_{d}'R^{-1}Z_{d}(Z_{d}'R^{-1}Z_{d} + P^{-1})^{-1})Z_{d}'R^{-1}Z_{d}$$
$$(I - Z_{d}'R^{-1}Z_{d}(Z_{d}'R^{-1}Z_{d} + P^{-1})^{-1})Z_{d}'R^{-1}y_{d}$$

where subscript *d* refers to the bull dam, and a^{dd} is the diagonal element of dam *d* in the inverse of the relationship matrix. After completing all necessary absorptions, MME pertinent to the bull and his parents' additive genetic effects can be written in a general form:

$$\begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \begin{bmatrix} \hat{s} \\ \hat{d} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} RHS_s \\ RHS_d \\ RHS_b \end{bmatrix}$$
[1]

Solving [1] gives:

$$\begin{bmatrix} \hat{s} \\ \hat{d} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix}^{-1} \begin{bmatrix} RHS_s \\ RHS_d \\ RHS_b \end{bmatrix} = \begin{bmatrix} C_{ss} & C_{sd} & C_{sb} \\ C_{ds} & C_{dd} & C_{db} \\ C_{bs} & C_{bd} & C_{bb} \end{bmatrix} \begin{bmatrix} RHS_s \\ RHS_d \\ RHS_b \end{bmatrix} = \begin{bmatrix} C_{ssdd} & C_{sd-b} \\ C_{b-sd} & C_{bb} \end{bmatrix} \begin{bmatrix} RHS_{sd} \\ RHS_b \end{bmatrix}$$

It is assumed that cows are randomly mated and not related to the test bull. An average reliability for bull mate proofs is set to $REL_{mate} = 0.40$ for test day milk yield at the time their daughters start lactation. Moreover, it is assumed that daughters' test day records yhave been corrected for all the fixed effects β , which are assumed to be estimated without errors, as well as for bull mate EBV. The adjusted test day records are analogue to yield deviations for the case of repeated records. The adjustment for bull mate EBV leads to a reduction genetic in variance by $0.25 REL_{mate} G_{o}$ (VanRaden and Wiggans, 1991).

For the derivation, assumptions on the reliabilities of bull parents are to be made. At

the time the first crop of daughters starts their lactation, reliability of bull sire proofs usually has already been very high, and the high reliability is simulated with $n_{ds} = 1000$ daughters with three complete lactations, composed of eight tests each. Because it is difficult to make a reasonably good estimate of the reliability of bull dams' proofs as it is the case for bull sires, a lower and upper limit are set up instead. In the case of lower limit (Worst Case Scenario), the bull dam is assumed to have contributed only three complete lactations of her own to her BV estimation, whereas in the case of upper limit (Best Case Scenario) the bull dam is assumed to have many offspring including sons and daughters and the highest reliability of dam proofs is simulated with $n_{dd} = 1000$ daughters with three complete lactations each.

Deriving weights for Best Case Scenario: Equations [1] for Best Case Scenario are:

$$\begin{bmatrix} a^{ss}G_{o}^{-1} - \sum_{j=1}^{n_{ds}} G_{o}^{-1}C_{j}^{11}G_{o}^{-1} & 0.5G_{o}^{-1} & -G_{o}^{-1} \\ 0.5G_{o}^{-1} & a^{dd}G_{o}^{-1} - \sum_{k=1}^{n_{dd}} G_{o}^{-1}C_{k}^{11}G_{o}^{-1} & -G_{o}^{-1} \\ -G_{o}^{-1} & -G_{o}^{-1} & a^{bb}G_{o}^{-1} - \sum_{i=1}^{n_{d}} G_{o}^{-1}C_{i}^{11}G_{o}^{-1} \end{bmatrix}$$

$$*\begin{bmatrix} \hat{s} \\ \hat{d} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} RHS_{s} \\ RHS_{d} \\ \sum_{i=1}^{n_{d}} G_{o}^{-1}(C_{i}^{11} + C_{i}^{12})Z_{i}'R^{-1}y_{i} \end{bmatrix}$$

$$[2]$$

where a^{ss} is the diagonal element of the sire in the inverse of the relationship matrix.

replacing RHS with parental EBV lead to the equations for bull's additive genetic effects:

Further absorbing parental equations and

$$\hat{b} = (C_{bb} - C_{b-sd}C_{ssdd}^{-1}C_{sd-b})\sum_{i=1}^{n_d} G_o^{-1}(C_i^{11} + C_i^{12})Z_i'R^{-1}y_i + C_{b-sd}C_{ssdd}^{-1}q$$
[3]

where $q = [\hat{s}' \hat{d}']'$.

Deriving weights for Worst Case Scenario Equations [1] for Worst Case Scenario are:

$$\begin{bmatrix} a^{ss}G_{o}^{-1} - \sum_{j=1}^{n_{ds}}G_{o}^{-1}C_{j}^{1I}G_{o}^{-1} & 0.5G_{o}^{-1} & -G_{o}^{-1} \\ 0.5G_{o}^{-1} & a^{dd}G_{o}^{-1} + (I - Z_{d}'R^{-1}Z_{d}(Z_{d}'R^{-1}Z_{d} + P^{-1})^{-1})Z_{d}'R^{-1}Z_{d} & -G_{o}^{-1} \\ -G_{o}^{-1} & -G_{o}^{-1} & a^{bb}G_{o}^{-1} - \sum_{i=1}^{n_{d}}G_{o}^{-1}C_{i}^{1I}G_{o}^{-1} \end{bmatrix}$$

$$*\begin{bmatrix} \hat{s} \\ \hat{d} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} RHS_{s} \\ (I - Z_{d}'R^{-1}Z_{d}(Z_{d}'R^{-1}Z_{d} + P^{-1})^{-1})Z_{d}'R^{-1}y_{d} \\ \sum_{i=1}^{n_{d}}G_{o}^{-1}(C_{i}^{11} + C_{i}^{12})Z_{i}'R^{-1}y_{i} \end{bmatrix}$$

$$[4]$$

As above, parental equations are absorbed and sire RHS are replaced with

his EBV. The final equations for bull's additive genetic effects are:

$$\hat{b} = C_{bs} C_{ss}^{-1} \hat{s} + (C_{bd} - C_{bs} C_{ss}^{-1} C_{sd}) (I - Z_d' R^{-1} Z_d (Z_d' R^{-1} Z_d + P^{-1})^{-1}) Z_d' R^{-1} y_d + (C_{bb} - C_{bs} C_{ss}^{-1} C_{sb}) \sum_{i=1}^{n_d} G_o^{-1} (C_i^{11} + C_i^{12}) Z_i' R^{-1} y_i$$
[5]

Weights on daughter and pedigree information can be computed directly using equations [3] and [5] for *Best* and *Worst Case Scenarios*, respectively. However, these weights are difficult to interpret and to compare across scenarios, therefore they are to be standardised. For *Best Case Scenario*, equations [3] can also be written in an equivalent form:

$$\hat{b} = (Q_{bb} + Q_{sd}) \{ (Q_{bb} + Q_{sd})^{-1} Q_{bb} \overline{y}_{b} + (Q_{bb} + Q_{sd})^{-1} Q_{sd} q \}$$
[6]

where
$$Q_{bb} = (C_{bb} - C_{b \cdot sd}C_{ssdd}^{-1}C_{sd \cdot b}) \sum_{i=1}^{n_d} G_o^{-1} (C_i^{11} + C_i^{12}) Z_i' R^{-1} Z_i, Q_{sd} = C_{b \cdot sd}C_{ssdd}^{-1},$$

 $(Q_{bb} + Q_{sd})^{-1} Q_{bb}$ are standard weights on daughter information,
 $(Q_{bb} + Q_{sd})^{-1} Q_{sd}$ are standard weights on pedigree information, and
 \overline{y}_b represents average of daughters' adjusted test day records by lactation.

Similarly, standard weights are derived for Worst Case Scenario:

$$\hat{b} = (Q_{bb} + Q_{ss} + Q_{dd})$$

$$\{ (Q_{bb} + Q_{ss} + Q_{dd})^{-1} Q_{ss} \hat{s} + (Q_{bb} + Q_{ss} + Q_{dd})^{-1} Q_{dd} \overline{y}_{d} + (Q_{bb} + Q_{ss} + Q_{dd})^{-1} Q_{bb} \overline{y}_{b} \}$$
[7]

where
$$Q_{ss} = C_{bs}C_{ss}^{-1}$$
,
 $Q_{dd} = (C_{bd} - C_{bs}C_{ss}^{-1}C_{sd})(I - Z_d'R^{-1}Z_d(Z_d'R^{-1}Z_d + P^{-1})^{-1})Z_d'R^{-1}Z_d$,
 $Q_{bb} = (C_{bb} - C_{bs}C_{ss}^{-1}C_{sb})\sum_{i=1}^{n_d} G_o^{-1}(C_i^{11} + C_i^{12})Z_i'R^{-1}Z_i$,
 $(Q_{bb} + Q_{ss} + Q_{dd})^{-1}Q_{ss}$ are standard weights on bull sire proofs,
 $(Q_{bb} + Q_{ss} + Q_{dd})^{-1}Q_{dd}$ are standard weights on bull dam records,
 $(Q_{bb} + Q_{ss} + Q_{dd})^{-1}Q_{bb}$ are standard weights on daughter information, and
 \overline{y}_d represents average of dam's adjusted records by lactation.

The same additive genetic, permanent environmental and residual (co)variances are used in this study for deriving the weights as for routine genetic evaluations of test day milk yield trait (Reents et al. 1998). Computer programs in Maple V (Monagan et al. 1998) were written for conducting this research.

Results and Discussion

Characteristics of standardised weights on daughter and pedigree information: For estimating BV of a given lactation, standard weights of all sources of information: daughter records and pedigree index of current lactation, daughter records and pedigree index from preceding or following lactations, sum to 100. Pedigree index and daughter records of preceding or following lactation have equal weights for estimating BV of the current lactation, but pedigree indices with a negative weight. Thus, only Mendelian sampling component of preceding or following lactation contributes to the estimation of BV of a given lactation. Weights on daughter records and pedigree index for estimating BV of the same lactation sum to 100 as well. Due to different genetic variances of lactations, weights on first lactation records for estimating later lactation EBV are greater than weights on later lactation records for estimating first lactation EBV.

Standard weights for Best and Worst Case Tables 1 and 2 show standard Scenarios: weights on daughter information and pedigree index (in bold) for estimating BVs of test bulls Best and Worst Case Scenarios. in respectively. In the two tables each row represents a lactation BV, and each column the information of a lactation contributing to the estimation of lactation BV. Some typical cases with different daughter information are explained below in detail.

All daughters with three complete lactations *each:* In the case of 100 daughters with three complete lactations each, the ratio of weight on daughter records to weight on pedigree index for estimating combined BV of the test bull is 95:5 under Best Case Scenario, which is the average of weight ratios for the three lactation EBV 92:8, 98:2 and 96:4, based on the definition of combined lactation EBV (Reents et al. 1998). In this case EBV of the test bull are predominantly determined by his daughters rather than by parental information. For the first lactation EBV, the weights on first lactation daughter records, first lactation pedigree index, second lactation daughter records and third lactation daughter records are respectively 84, 16, 4, and 4 under Best Case Scenario. Additionally, the weights on second and third lactation pedigree index are -4 and -4 for estimating first lactation BV, due to the equal weighting factors on daughter records and pedigree index of preceding or following lactation. Summing up the weights on daughter records or pedigree indices over all lactations gives 92:8 for first lactation EBV under Best Case Scenario. Comparing the two tables, it is obvious that different reliabilities of bull dam EBV lead to the differences in weights on daughter and pedigree information for estimating both individual lactation and combined BV for test bulls, although the differences are of relatively small magnitude. In general, daughter information from current lactation makes a larger contribution to the estimation of a lactation BV for the bull in *Worst Case Scenario* than in *Best Case Scenario*.

Daughters with first lactation records only Since no later lactation records are available. first lactation BV is estimated only with first lactation daughter and pedigree information. First lactation Mendelian sampling component and second (third) lactation pedigree index contribute to the estimation of second (third) lactation BV. The ratio of weights on first lactation Mendelian sampling component and later lactation pedigree index does not change much regardless of the number of daughters or records. This ratio approaches its maximum value, which is equal to the ratio of genetic covariance to genetic variance of the first lactation, as more daughter information from first lactation is accumulated. As a result of relatively high heritability of first lactation test day milk yield, the ratio of weights on first lactation daughter information and pedigree index for estimating first lactation BV increases little when the number of daughters increases from 50 to 100. Similarly due to the high repeatability, the contribution of daughter information to first lactation BV of the test bull stays almost unchanged from 3 to 8 tests. When 100 daughters all have completed their first lactations, weight on daughter records to weight on pedigree index over all three lactations is only 59:41 under Best Case Scenario, indicating that bull parents still have a substantial impact on bull's overall EBV even for bull with 100 complete first lactations. As expected, first lactation daughter information makes larger contribution to first lactation EBV of the test bull in Worst Case Scenario than in Best Case Scenario. And so is the Mendelian sampling component of first lactation daughter records to second or third lactation EBV, though the difference in the contributions between the two scenarios is marginal.

Summary

A method was presented to derive standardised weights on daughter records and pedigree index for estimating BV of progeny tested bulls using a test day model that treats lactations as genetically and environmentally different traits. Pedigree index from other lactation does not make any contribution to the estimation of BV of a given lactation, only Mendelian sampling component of a lactation affects EBV of correlated lactation. The impact of Mendelian sampling component from other lactation on the estimation of BV of a given lactation is reduced drastically, as soon as daughter records of the lactation are available. After many daughters have undergone three milk tests in the third lactation, weight on daughter information for estimating overall BV of test bulls reaches its maximum, given the genetic parameters used. The derivation method is applicable to cows or traits with single records and can also be used for investigating the effects of e.g.

overestimated bull dam EBV or poor lactation persistency, on stability of bull EBV.

Literature Cited

- Monagan, M.B., Geddes, K. O., Heal, K.M., Labahn, G. & Vorkoetter, S.M. 1998. Maple V programming guide. Springer Verlag.
- Reents, R., Jamrozik, J., Schaeffer, L.R. & Dekkers, J.C.M. 1995. Estimation of genetic parameters for test day records of somatic cell scores. J. Dairy Sci. 78, 2847-2857.
- Reents, R., Dopp, L., Schmutz M. & Reinhardt F. 1998. Impact of application of a test day model to dairy production traits on genetic evaluations of cows. Interbull Bulletin No.17, 49-54.
- VanRaden, P.M. & Wiggans, G.R. 1991. Derivation, calculation, and use of national animal model information. J. Dairy. Sci. 74, 2737-2746.

Table 1. Standard weights on daughter and pedigree information for estimating breeding
values of test bulls under Best Case Scenario (each row/column represents a lactation
EBV/information from a lactation)

No. of daughters	No. of tests per lactation	Standard v ped	veights on da igree informa	By lactation	Over all lactations	
50	3/0/0	85:15	<u> </u>		85:15	
		97	100		49: 51	
		95		100	49: 51	56: 44
100	8 / 0 / 0	93: 7			93: 7	
		106	100		51: 49	
		104		100	51: 49	59: 41
100	8 / 8 / 0	84: 16	8		92: 8	
		21	76: 24		97: 3	
		17	78	100	49: 51	72: 28
70	8 / 8 / 0	84: 16	6	2	92: 8	
30	8 / 8 / 3	20	69: 31	8	97: 3	
		13	50	32: 68	95: 5	95: 5
100	8 / 8 / 8	84: 16	4	4	92: 8	
		19	61: 39	18	98: 2	
		9	17	70: 30	96: 4	95: 5

For estimating first lactation EBV, weight on first lactation daughter records and on first lactation pedigree index are 84 and 16, respectively.

For estimating first lactation EBV, weight on second lactation daughter records (second lactation pedigree index) is 4 (-4).

For estimating third lactation EBV, weight on first lactation daughter records (first lactation pedigree index) is 9 (-9).

Table 2. Standard weights on daughter and pedigree information for estimating breeding values of test bulls under *Worst Case Scenario* (each row/column representing a lactation EBV/information from a lactation)

No. of	No. of tests	Standard v	veights on da	By	Over all	
daughters	per lactation	ped	igree informa	lactation	lactations	
50	3 / 0 / 0	91: 9			91: 9	
		101	100		50: 50	
		98		100	49: 51	58: 42
100	8 / 0 / 0	96: 4			96: 4	
		107	100		52: 48	
		103		100	51: 49	60: 40
100	8 / 8 / 0	92: 8	4		96: 4	
		11	87: 13		98: 2	
		47	59	100	51: 49	74: 26
70	8 / 8 / 0	92: 8	3	1	96: 4	
30	8 / 8 / 3	11	81: 19	7	98: 2	
		5	43	49: 51	97: 3	97: 3
100	8 / 8 / 8	92: 8	2	2	96: 4	
		10	77: 23	11	98: 2	
		3	11	83: 17	97: 3	97: 3