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

Z. Liu, F. Reinhardt, and R. Reents
VIT, Heideweg 1, D-27280 Verden/Aller, Germany

INTRODUCTION
An effective daughter contribution method (EDC, Fikse and Banos, 2001) was extended for
reliability approximation for general multiple trait models (Liu et al., 2001a) and an
application of the multiple trait EDC (MTEDC) was outlined for a random regression test day
model (RRTDM) for genetic evaluation in dairy cattle (Liu et al., 2001b). The objectives of
the current study were to investigate the efficiency of MTEDC approach for a very large
population, and to study the accuracy of MTEDC for two groups of animals.

MATERIAL AND METHODS
Genetic evaluation model. Test day yields on a 24-hour daily basis from first three lactations
are analysed with a RRTDM (Liu et al., 2001b) separately for milk, fat, protein yield and
somatic cell scores. The model includes fixed herd-test-date-parity-milking-frequency (HTD)
effects, fixed lactation curve effects, random permanent environmental and additive genetic
effects, and error effects. The third-order Legendre polynomials are used for modelling both
random effects and the Wilmink function for modelling fixed lactation curves.

Field data set. Test day data from August 2001 national genetic evaluation for German
Holstein and Red breeds were used for testing the efficiency of MTEDC. Table 1 presents the
size of data set after all edits. For the test run the total number of equations to be solved
amounted to approximately 227 millions per trait.

Table 1. Size of the data set used in August 2001 test run of genetic evaluation for
German Holstein and Red breeds.

<table>
<thead>
<tr>
<th></th>
<th>Cows with records</th>
<th>Animals in pedigree</th>
<th>Lactations</th>
<th>HTD</th>
<th>Test day records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>9,251,847</td>
<td>13,979,772</td>
<td>18,627,806</td>
<td>13,114,434</td>
<td>146,908,758</td>
</tr>
</tbody>
</table>

Simulated data sets. A thorough study of the accuracy of MTEDC requires extensive
programming work, therefore, two scenarios where true reliability values can be easily
obtained were considered: (i) lactation progress of a typical cow in monthly milk testing
programme was simulated, where only her own test day records contribute to her estimated
breeding values (EBV); (ii) a progeny-tested bull was assumed to have no known parents, and
only his daughters performance records made contribution to his EBV. Further it was assumed

Session 20. Prediction of breeding values
that all of his daughters are in identical lactation stages and their dams are missing. For the cow or the daughters of the progeny-tested bull, a new test day record was added every 30 days, starting from days in milk 10. A total of 10 tests were generated for each of the first three lactations. The analysed trait is test day milk yield of first three lactations and parameters of the trait were obtained from Liu et al. (2000).

MTEDC applied to the field data set. Following the procedure of the daughter equivalents method (VanRaden and Wiggans, 1991), MTEDC was applied to approximate reliability values of EBV of random regression coefficients (RRC). Reliability values of EBV on a 305-day lactation basis were then calculated using the reliabilities of RRC (Liu et al., 2001b).

Computing true reliability. For the simulated cow, true reliability matrix of RRC is

\[ \Phi_{\text{true}}^{\text{cow}} = I - G_0^{-1}(B_c + G_0^{-1})^{-1} \]

with \( G \) being genetic (co)variance matrix of RRC and \( B \) being effective test day data information after absorbing all but genetic effects. True reliability matrix of RRC for the simulated bull is

\[ \Phi_{\text{true}}^{\text{bull}} = I - G_0^{-1}B_0^{-1} \]

where \( B_0 = G_0^{-1}[(1 + \frac{2}{3})I - \frac{4}{5} \sum_{i=1}^{n}((B_i + \frac{4}{3}G_0^{-1}G_0^{-1}))] \) with \( n \) being the number of daughters and \( B_i \) being effective test day data information of daughter \( i \).

Approximating reliability based on the effective number of progeny method. In addition to the comparison of true reliability values and approximated ones using MTEDC, the Effective Number of Progeny (ENP, Jamrozik et al., 2000) was also applied to the simulated data sets. Heritability values of individual lactation as well as combined lactation breeding values were derived for the application of ENP method.

RESULTS AND DISCUSSION
Application of MTEDC to the field data set. Fortran 90 programmes were developed for approximating reliability values of EBV for the RRTDM and were tested on a HP Unix computer HP9000/L2000 using the field data set. Half-stored EDC matrices for all animals were kept in RAM in double precision. The total RAM usage was 4.9Gb which is equivalent to 90% RAM consumption of the iteration programme with a lactation based iteration scheme (Liu et al., 2001b). Table 2 shows CPU usage for each step of MTEDC applied to the data set for each trait. It can be seen that reliability approximation using MTEDC requires less RAM and much less CPU than its corresponding iteration programme (Liu et al., 2001b). The application of MTEDC proved that it is feasible to implement MTEDC to approximate reliability values of EBV for a very large population like the German Holstein and Red breeds.

Results for the simulated data sets. Maple 6 programmes were developed to compute reliability values for the bull or the cow. For the cow identical reliability values of individual and combined lactation EBV were obtained using all the methods: the true reliability method, MTEDC and ENP. For the progeny-tested bull, MTEDC and the true reliability methods give
equal reliability values. It can be seen in Figure 1 that the increase in reliability of each individual lactation EBV is slowed down when daughters have no more test day records available for the lactation. Compared to the true reliability values, it can be seen in Figure 1 that the method ENP clearly provides biased reliability values. ENP severely overestimated reliability values for the bull when his daughters had first lactation test day records only. When all daughters completed their first three lactations, the reliability estimates by ENP were slightly lower than their true values.

Table 2. CPU consumption for approximating reliabilities using MTEDC applied to the field data set.

<table>
<thead>
<tr>
<th>Step of computing</th>
<th>Data contribution</th>
<th>Progeny contribution</th>
<th>Parental contribution</th>
<th>Reliability of 305-day EBV</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU usage (minutes)</td>
<td>17.4</td>
<td>114.7</td>
<td>154.7</td>
<td>18.5</td>
<td>305.5</td>
</tr>
<tr>
<td>Equivalent number of rounds of iteration</td>
<td>1.4</td>
<td>9.2</td>
<td>12.4</td>
<td>1.5</td>
<td>24.4</td>
</tr>
</tbody>
</table>

Discussion. MTEDC was demonstrated here to be efficient and feasible for approximating reliability values of EBV for very large populations. Though MTEDC was tested here using a single trait multiple lactation RRTDM, it was also applied to a multiple trait multiple lactation RRTDM for the Dutch dairy cattle population (De Roos, 2001). Calculating parental contribution to an animal requires re-computation of the animal’s contribution to its parents that was done in the previous step, and this can be avoided for higher run-time performance if animal’s contribution to its parents can be stored in the step of progeny contribution calculation and then be read for computing parental contribution. ENP caused severe upward bias in reliability values for the bull when his daughters have missing or short lactations, though no bias was found for the cow with only own performance records. It seems that the way that progeny information is converted to the progeny equivalents in ENP is not suitable for multiple trait models, although ENP works well for single trait models.

No special handling of missing traits is needed in MTEDC, but MTEDC makes approximation in the calculation of parental contribution as all existing reliability approximation methods. Also for an inbred progeny, its contribution to parents is not computed exactly. The magnitude of these biases must be examined via simulation studies. A problem in application of MTEDC was observed (De Roos, 2001) for the calculation of progeny contribution: at the time of adjusting the contribution of a progeny to an animal for the reliability of animal’s mate the mate may not yet have all its progeny processed. This problem could be partially solved by using reliability values from previous genetic evaluation as starting values for the mates, however, this bias has limited impact due to the factor 0.25 for mate’s reliability matrix. Although MTEDC has been shown to be fairly accurate for dairy cattle, it should also be applicable to other livestock species, e.g. swine or beef cattle. MTEDC can also be used for approximating reliability values for multiple trait models with correlated genetic effects by
treating the correlated genetic effects as a unit. In this case the definition of one EDC (Liu et al., 2001a) changes accordingly.

Figure 1. Reliability values of single and combined lactation EBV using the true reliability and MTEDC methods (solid lines) and biases in reliability using the ENP method (dashed lines) for the bull

REFERENCES