

A Multi-Trait MACE Model for International Bull Comparison Based on Daughter Yield Deviations

Z. Liu, F. Reinhardt and R. Reents
VIT, Heideweg 1, D-27283 Verden, Germany

Abstract

A multi-trait MACE model is presented for international bull comparison based on daughter yield deviations. This model allows a variable number of correlated traits per country. For countries using multiple trait model in national genetic evaluation, a vector of daughter yield deviations and corresponding effective daughter contribution matrix are needed for each bull having daughter performance information. An approximate REML algorithm is developed for estimating across country genetic correlations. Prediction error (co)variances of Mendelian sampling estimates are approximated based on the multi-trait effective daughter contribution method. Due to the large mixed model equations, a computing algorithm is presented to solve the equation system efficiently with preconditioned conjugate gradients. The model is to be validated with the data of a number of traits from German national genetic evaluation. The traits are to be evaluated as if they were the same biological traits from different countries in international genetic evaluation.

1. Introduction

The multiple across country evaluation (MACE; Schaeffer, 1994) method is used for international bull comparisons. Estimated breeding values (EBV) from each country are deregressed to obtain an equivalent value of daughter yield deviations (DYD) for bulls that have daughters with records. One single EBV per bull is permitted for each country in current international genetic evaluation. As more and more countries have upgraded their national genetic evaluation systems to a multiple trait model, e.g. a multiple lactation random regression test day model (RRTDM), differences between the models for national and international evaluations become increasingly evident. In order to optimise the genetic evaluation models for both national and international evaluations, Sullivan and Wilton (2001) proposed a multiple trait MACE with a variable number of traits per country. This model extended the current single trait MACE to multiple lactations/traits for countries using a multiple trait model in national genetic evaluation. Schaeffer (2001) developed a multiple trait deregression method for MACE evaluation.

In parallel to the development involving deregression, Ducrocq *et al.* (2001) suggested to use corrected records for a two-step multiple

trait genetic evaluation. Indeed, this approach can also be applied to international genetic evaluations. Yield deviations (YD) of cows and DYD of bulls are, in fact, corrected records. For single trait models, VanRaden and Wiggans (1991) developed formulae for the calculation of YD and DYD. Liu *et al.* (2004) extended the YD and DYD methods to multiple trait models, including RRTDM. Also, Liu *et al.* (2004) developed a method for approximating reliabilities of EBV under general multiple trait models, which can also be utilized to approximate effective daughter contribution (EDC) associated with bulls' DYD. The objectives of this study are 1) to present a multiple trait MACE model, based on DYD and EDC, for international evaluation of production traits, 2) to apply the preconditioned conjugate gradients algorithm to solving the equation system of the model, and 3) to develop an approximate REML method for estimating across country genetic correlations.

2. Materials and Methods

2.1. A multi-trait MACE model based on DYD

An ideal genetic model of international genetic evaluation for a country should be identical to the genetic model of its national genetic evaluation, as far as additive genetic effects

concerned. A statistical model for evaluating DYD of bulls is assumed to be:

$$\mathbf{y}_{ij} = \mathbf{f}_j + \mathbf{a}_{ij} + \mathbf{e}_{ij} \quad [1]$$

where \mathbf{y}_{ij} is a vector of DYD for bull i ($i = 1, 2, \dots, n$) in country j ($j = 1, 2, \dots, m$), \mathbf{f}_j is a vector of general mean effects of country j , \mathbf{a}_{ij} is a vector of additive genetic effects of

bull i in country j , and \mathbf{e}_{ij} is a vector of error effects. The model [1] allows countries to have multiple correlated traits per bull, e.g. DYD regression coefficients of first three lactations. In case that a country applies a single trait model for its national evaluation, all corresponding vectors in model [1] become scalar. (Co)variance matrix of genetic effects of the m countries and its inverse are defined as:

$$\mathbf{G}_0 = \begin{bmatrix} \mathbf{G}_{0_{11}} & \mathbf{G}_{0_{12}} & \cdots & \mathbf{G}_{0_{1m}} \\ & \mathbf{G}_{0_{22}} & \cdots & \mathbf{G}_{0_{2m}} \\ & & \ddots & \vdots \\ \text{symm.} & & & \mathbf{G}_{0_{mm}} \end{bmatrix}, \text{ and } \mathbf{G}_0^{-1} = \begin{bmatrix} \mathbf{G}_0^{11} & \mathbf{G}_0^{12} & \cdots & \mathbf{G}_0^{1m} \\ & \mathbf{G}_0^{22} & \cdots & \mathbf{G}_0^{2m} \\ & & \ddots & \vdots \\ \text{symm.} & & & \mathbf{G}_0^{mm} \end{bmatrix}, \quad [2]$$

where $\mathbf{G}_{0_{jj}}$ is original genetic (co)variance matrix of country j , and $\mathbf{G}_{0_{jk}}$ is genetic covariance matrix between country j and k . When country j uses a single trait model for national evaluation, $\mathbf{G}_{0_{jj}}$ is a scalar and all corresponding off-diagonal blocks, $\mathbf{G}_{0_{jk}}$ ($j \neq k$), become vector or scalar, depending the number of traits in other countries. As usually assumed in MACE, DYD of a bull from different countries are assumed to be residually uncorrelated. The inverse of error (co)variance matrix of DYD of bull i in country j is:

$$[\text{Var}(\mathbf{e}_{ij})]^{-1} = \Psi_{ij} \quad [3]$$

$$\begin{bmatrix} \Psi_{i1} & & & \mathbf{0} \\ & \Psi_{i2} & & \\ & & \ddots & \\ \mathbf{0} & & & \Psi_{im} \end{bmatrix} + a^{ii} \begin{bmatrix} \mathbf{G}_0^{11} & \mathbf{G}_0^{12} & \cdots & \mathbf{G}_0^{1m} \\ & \mathbf{G}_0^{22} & \cdots & \mathbf{G}_0^{2m} \\ & & \ddots & \vdots \\ \text{symm.} & & & \mathbf{G}_0^{mm} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{a}}_{i1} \\ \hat{\mathbf{a}}_{i2} \\ \vdots \\ \hat{\mathbf{a}}_{im} \end{bmatrix} + \begin{bmatrix} \Psi_{i1} & & & \mathbf{0} \\ & \Psi_{i2} & & \\ & & \ddots & \\ \mathbf{0} & & & \Psi_{im} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_1 \\ \hat{\mathbf{f}}_2 \\ \vdots \\ \hat{\mathbf{f}}_m \end{bmatrix} = \begin{bmatrix} \Delta_{i1} \\ \Delta_{i2} \\ \vdots \\ \Delta_{im} \end{bmatrix} \quad [4]$$

where a^{ii} is diagonal element of bull i in the inverse of numerator relationship matrix, and Δ_{ij} represents right-hand-sides of bull i in country j . Note that Δ_{ij} is not dependent on

where Ψ_{ij} is EDC matrix for DYD of bull i in country j , converted from reliability contributed by daughters' records in j -th country. Ψ_{ij} is defined for multiple trait models with formula 23 in the paper by Liu *et al.* (2004).

2.2. Mixed model equations of the multi-trait MACE model

Mixed model equations (MME) of model [1] consist of genetic effects of bulls and fixed effects of country means. Ignoring pedigree contribution, a bull i has following equations:

across country correlations and is a function of country specific information only. Formula 24 in the paper by Liu *et al.* (2004) can be used to calculate Δ_{ij} . Equations for the fixed effects of country means in model [1] are:

$$\begin{bmatrix} \Psi_{f_1} & & & \mathbf{0} \\ & \Psi_{f_2} & & \\ & & \ddots & \\ \mathbf{0} & & & \Psi_{f_m} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{f}}_1 \\ \hat{\mathbf{f}}_2 \\ \vdots \\ \hat{\mathbf{f}}_m \end{bmatrix} + \begin{bmatrix} \Psi_{11} & & & \mathbf{0} \\ & \Psi_{12} & & \\ & & \ddots & \\ \mathbf{0} & & & \Psi_{nm} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{a}}_1 \\ \vdots \\ \hat{\mathbf{a}}_n \end{bmatrix} = \begin{bmatrix} \sum_i \Delta_{i1} \\ \sum_i \Delta_{i2} \\ \vdots \\ \sum_i \Delta_{im} \end{bmatrix} \quad [5]$$

where $\hat{\mathbf{a}}_i' = [\hat{\mathbf{a}}_{i1}' \ \hat{\mathbf{a}}_{i2}' \ \dots \ \hat{\mathbf{a}}_{im}']$ represents EBV of bull i in all m countries, and

$\Psi_{f_j} = \sum_{i=1}^{n_j} \Psi_{ij}$ with n_j being the number of bulls with DYD in country j .

Ancestors of the bulls with DYD are traced back as far as possible based on sire and dam, instead of sire and maternal grandsire as it is now in current international genetic evaluations. Phantom groups should be defined across countries in the same way for both parameter estimation and genetic evaluation.

2.3. Solving the equation system

Under the multi-trait MACE model [1], the original (co)variance parameters from all countries are kept constant. Therefore, least squares parts of left-hand-sides of the MME, e.g. Ψ_{ij} for bull i in country j , are not dependent on the correlations between countries. Because bull effects in a country are nested within the country mean effect, off-diagonal block between a bull and the country is equal to his least squares part of left-hand-side of the coefficient matrix, i.e. Ψ_{ij} . Therefore, no off-diagonal blocks between genetic and fixed effects need to be set up explicitly. The equation system can be solved using pre-conditioned conjugate gradients algorithm (PCG, Strandén and Lidauer, 1999) incorporated with the iteration on data technique (Schaeffer and Kennedy, 1986).

Four vectors with size of the MME are required for using PCG: solution vector $[\mathbf{f}' \ \mathbf{a}']$, search direction vector \mathbf{z} , residual vector \mathbf{r} , and work vector \mathbf{w} . The most costly step of solving equations with PCG algorithm involves a multiplication of the coefficient matrix of the MME with search direction

vector \mathbf{z} , which consists of accumulation of data contribution and accumulation of pedigree contribution.

2.3.1. Calculation of data contribution

For the multiplication of the coefficient matrix with direction vector, files containing EDC matrices of bulls are processed on a country by country basis to accumulate data contribution. Diagonal block, e.g. Ψ_{f_j} for country j , can be computed by summing up diagonal blocks of bulls in the j -th country, prior to iteration process.

At round t of PCG iteration, the data are processed in the order of bulls within country to accumulate information for the fixed effects:

$$\mathbf{w}_{f_j} = \mathbf{w}_{f_j} + \Psi_{f_j} \hat{\mathbf{z}}_{f_j}^{[t-1]} + \sum_{i=1}^{n_j} \Psi_{ij} \hat{\mathbf{z}}_{ij}^{[t-1]} \quad [6]$$

where \mathbf{w}_{f_j} is a sub-vector of \mathbf{w} corresponding to the fixed effects \mathbf{f}_j , $\hat{\mathbf{z}}_{f_j}^{[t-1]}$ corresponds to the fixed effects \mathbf{f}_j at iteration round $t-1$, and $\hat{\mathbf{z}}_{ij}^{[t-1]}$ is a sub-vector of search direction vector \mathbf{z} corresponding to genetic effects of bull i in country j at iteration round $t-1$. At the same round of iteration, contribution of the i -th bull's data in country j is calculated with:

$$\mathbf{w}_{ij} = \Psi_{ij} (\hat{\mathbf{z}}_{ij}^{[t-1]} + \hat{\mathbf{z}}_{f_j}^{[t-1]}). \quad [7]$$

2.3.2. Calculation of pedigree contribution

For the multiplication of the coefficient matrix and direction vector, pedigree file must be read sequentially to accumulate pedigree contribution. For animal i , its pedigree contribution to the multiplication is:

$$\mathbf{w}_i = \mathbf{w}_i + a^{ii} \mathbf{G}_0^{-1} \hat{\mathbf{z}}_i^{[t-1]} - \frac{1}{2} d_i \mathbf{G}_0^{-1} (\hat{\mathbf{z}}_s^{[t-1]} + \hat{\mathbf{z}}_d^{[t-1]}), \quad [8]$$

$$\mathbf{w}_s = \mathbf{w}_s - \frac{1}{2} d_i \mathbf{G}_0^{-1} \hat{\mathbf{z}}_i^{[t-1]} + \frac{1}{4} d_i \mathbf{G}_0^{-1} \hat{\mathbf{z}}_d^{[t-1]},$$

and [9]

$$\mathbf{w}_d = \mathbf{w}_d - \frac{1}{2} d_i \mathbf{G}_0^{-1} \hat{\mathbf{z}}_i^{[t-1]} + \frac{1}{4} d_i \mathbf{G}_0^{-1} \hat{\mathbf{z}}_s^{[t-1]} \quad [10]$$

where subscripts s and d represent sire and dam of animal i , \mathbf{w}_i , \mathbf{w}_s , and \mathbf{w}_d are sub-vectors of \mathbf{w} for genetic effects of animal i and its sire s and dam d , respectively. d_i is diagonal element pertinent to animal i in the inverse of Mendelian sampling matrix \mathbf{D}^{-1} (Mrode, 1996, pp28). For genetic groups, only the term $a^{ii} \mathbf{G}_0^{-1} \hat{\mathbf{z}}_i^{[t-1]}$ of formula [8] is relevant, because neither parent of genetic groups is known.

The second most costly step of PCG is the multiplication of pre-conditioner matrix with residual vector \mathbf{r} . A simple pre-conditioner is defined as diagonal block of each effect in the model [1]. The multiplication of the pre-conditioner matrix with residual vector can be done by first processing the fixed effects and then the pedigree file.

As a result of the fixation of original country parameters, diagonal blocks for all fixed effects do not change throughout the whole parameter estimation process and solving of the equations. In contrast, diagonal blocks for genetic effects of bulls having DYD change in accordance with across country correlation estimates, but they must be kept fixed during the process of PCG solving of the equations. The diagonal blocks for the fixed effects and genetic effects of bulls with records are inverted and stored in RAM or disk for solving the equation system. For animals without own data, diagonal blocks are equal to $a^{ii} \mathbf{G}_0^{-1}$ with inverse $1/a^{ii} \mathbf{G}_0$. Convergence of the iteration process is assessed by logarithm of the sum of squared changes in solutions between current and previous rounds divided by sum of squared solutions in the current round of iteration.

2.4. Estimation of (co)variance components

Genetic (co)variances are estimated with the Expectation Maximization Restricted Maximum Likelihood (EM-REML) algorithm:

$$\hat{\sigma}_{g_{jk}} = [\hat{\mathbf{a}}_j' \mathbf{A}^{-1} \hat{\mathbf{a}}_k + \text{tr}(\mathbf{A}^{-1} \mathbf{C}^{jk})] / q \quad [11]$$

where $\hat{\sigma}_{g_{jk}}$ is estimated genetic covariance between countries j and k , $\hat{\mathbf{a}}_j$ is a vector of EBV of all animals in country j (animals sorted within trait), \mathbf{C}^{jk} is a submatrix of the inverse of coefficient matrix, corresponding to the country pair j and k , q is the total number of animals in pedigree. For countries having multiple correlated traits, j and k represent the traits within country.

As the coefficient matrix of MME is infeasible to invert for the large equation system, alternative formulation of the EM-REML formula is sought. The numerator relation matrix can be decomposed as:

$$\mathbf{A} = \mathbf{T} \mathbf{D} \mathbf{T}' \quad [12]$$

where matrix \mathbf{T} is a lower triangular matrix and its inverse \mathbf{T}^{-1} is also a lower triangular matrix with 1 in the diagonals and the only non-zero element, -0.5 , between an animal and its known parent. \mathbf{D} is a diagonal matrix for Mendelian sampling (Mrode, 1996). As Mendelian sampling effects are $\mathbf{m}_j = \mathbf{T}^{-1} \mathbf{a}_j$, the EM-REML formula [11] can be reformulated (Fikse *et al.*, 2003; Sullivan, 2004) as:

$$\hat{\sigma}_{g_{jk}} = \frac{[\hat{\mathbf{m}}_j' \mathbf{D}^{-1} \hat{\mathbf{m}}_k + \text{tr}(\mathbf{D}^{-1} \text{diag}(\mathbf{T}^{-1} \mathbf{C}^{jk} (\mathbf{T}^{-1})'))]}{q} \quad [13]$$

The first term of formula [13] is a simple summation of Mendelian sampling estimates over all animals:

$$\hat{\mathbf{m}}_j' \mathbf{D}^{-1} \hat{\mathbf{m}}_k = \sum_{i=1}^q d_i \hat{m}_{ij} \hat{m}_{ik} \quad [14]$$

where \hat{m}_{ij} and \hat{m}_{ik} are Mendelian sampling estimates of animal i in country (or within

country trait) j and k , respectively. For animal i , corresponding diagonal element of the matrix product $\mathbf{T}^{-1}\mathbf{C}^{jk}(\mathbf{T}^{-1})'$ is:

$$\{diag(\mathbf{T}^{-1}\mathbf{C}^{jk}(\mathbf{T}^{-1})')\}_i = \mathbf{t}'\mathbf{C}_i^{jk}\mathbf{t} \quad [15]$$

where $\mathbf{t}' = [1 \quad -\frac{1}{2} \quad -\frac{1}{2}]$, and \mathbf{C}_i^{jk} is a sub-matrix of the inverse of coefficient matrix for country (sub-trait) pair j and k corresponding to animal i and its sire s and dam d . The approximation of \mathbf{C}_i^{jk} is given in details later.

From the above formulae [13] and [15], it can be seen that, in addition to diagonal elements of the inverse of the coefficient matrix, off-diagonal elements between animal, sire and dam are also required for estimating genetic (co)variance components. Because obtaining the inverse of coefficient matrix is usually infeasible for large scale equation systems, multiple trait effective daughter contribution (MTEDC, Liu *et al.*, 2004) is used to provide approximate diagonal elements of the inverse matrix. An approximation method is presented here for calculating off-diagonal elements of the inverse matrix corresponding to each triplet: animal, sire and dam.

2.4.1. Approximating (co)variances of Mendelian sampling estimates

In order to approximate the (co)variances of Mendelian sampling estimates for each animal, following two assumptions are made. Firstly, an animal is the only common progeny of its parents, which is less restrictive for livestock species like dairy cattle. Secondly, there are no environmental covariances between any two of the triplet, which can be easily met for model [1] containing almost only additive genetic effects. Under these two assumptions all required elements of the submatrix \mathbf{C}_i^{jk} can be obtained from inverting the following matrix for each animal:

$$\mathbf{C}_{isd} = \begin{bmatrix} \Psi_{i-sd}^* & -\frac{1}{2}d_i\mathbf{G}_0^{-1} & -\frac{1}{2}d_i\mathbf{G}_0^{-1} \\ & \Psi_{s-i}^* & \frac{1}{4}d_i\mathbf{G}_0^{-1} \\ \text{symm.} & & \Psi_{d-i}^* \end{bmatrix}, \quad [16]$$

where Ψ_{i-sd}^* is EDC matrix of animal i after taking account of all sources of information, except parental contribution to itself, Ψ_{s-i}^* (Ψ_{d-i}^*) is EDC matrix of sire (dam) of animal i after taking account of all sources of information but the contribution of animal i to the sire (dam). The matrix \mathbf{C}_{isd} is inverted for animal i :

$$\mathbf{C}_{isd}^{-1} = \begin{bmatrix} \mathbf{C}_i^{ii} & \mathbf{C}_i^{is} & \mathbf{C}_i^{id} \\ & \mathbf{C}_i^{ss} & \mathbf{C}_i^{sd} \\ \text{symm.} & & \mathbf{C}_i^{dd} \end{bmatrix}. \quad [17]$$

Note that \mathbf{C}_i^{jk} is a 3x3 matrix and represents a sub-matrix of \mathbf{C}_{isd}^{-1} .

The EDC matrices Ψ_{i-sd}^* , Ψ_{s-i}^* and Ψ_{d-i}^* are expressed on animal basis and not on progeny basis. MTEDC (Liu *et al.*, 2004) can be used to efficiently obtain the EDC matrices for all animals. The effective information, denoted as $\mathbf{Z}'\mathbf{W}\mathbf{Z}$ in Liu *et al.* (2004), contributed by his own data of bull i , is obtained by absorbing the other effects \mathbf{f} :

$$\mathbf{Z}'\mathbf{W}\mathbf{Z}_i = diag\{\Psi_{i1}^*, \Psi_{i2}^*, \dots, \Psi_{im}^*\}, \quad [18]$$

$$\text{where: } \Psi_{ij}^* = \Psi_{ij}(\mathbf{I} - \Psi_{f_j}^{-1}\Psi_{ij}). \quad [19]$$

Animal's contribution to parents is accumulated by processing pedigree file from the youngest to oldest animal. For each animal in pedigree file, compute and store EDC matrices for its contribution to sire (dam), Ψ_{i2s} (Ψ_{i2d}). Then the pedigree file is processed in the reverse order to collect parental contribution to each animal, e.g. sire's contribution to the animal Ψ_{s2i} and dam's contribution to the animal Ψ_{d2i} . Let Ψ_{Ti} , Ψ_{Ts} and Ψ_{Td} represent total EDC for animal, sire and dam, respectively. For each animal, compute the following EDC matrices:

$$\Psi_{i-sd} = \Psi_{Ti} - \Psi_{s2i} - \Psi_{d2i} \quad [20]$$

$$\Psi_{s-i} = \Psi_{Ts} - \Psi_{i2s} \quad [21]$$

$$\Psi_{d-i} = \Psi_{Td} - \Psi_{i2d}, \quad [22]$$

the three EDC matrices are then converted, using formula 4 in Liu *et al.* (2004), to reliability matrices \mathfrak{R}_{i-sd} , \mathfrak{R}_{s-i} and \mathfrak{R}_{d-i} , respectively. Finally, the diagonal elements of \mathbf{C}_{isd} are obtained with:

$$\Psi_{i-sd}^* = [(\mathbf{I} - \mathfrak{R}_{i-sd})^{-1} - \mathbf{I}] \mathbf{G}_0^{-1} \quad [23]$$

$$\Psi_{s-i}^* = [(\mathbf{I} - \mathfrak{R}_{s-i})^{-1} - \mathbf{I}] \mathbf{G}_0^{-1} \quad [24]$$

$$\Psi_{d-i}^* = [(\mathbf{I} - \mathfrak{R}_{d-i})^{-1} - \mathbf{I}] \mathbf{G}_0^{-1}. \quad [25]$$

2.4.2 Procedure for estimating (co)variances across countries

The EM-REML algorithm [13], expressed in Mendelian sampling effects, can be combined with PCG algorithm to estimate (co)variance parameters of the multi-trait MACE model [1]. A procedure developed for the estimation of across country correlations comprises four steps.

Step 1: Solve the equation system of the multi-trait MACE model with PCG.

Step 1a: Accumulate data contribution using formulae [6] and [7],

Step 1b: Accumulate pedigree contribution using formulae [8] to [10],

Repeat *steps 1a* and *1b* until the pre-specified convergence criteria are met.

Step 2: Compute the quadratic term for a given covariance component with formula [14].

Step 3: Approximate prediction error (co)variances (PEC) of Mendelian sampling estimates with MTEDC and formulae [16] to [25].

Step 4: Estimate genetic (co)variances using the EM-REML formula [13] and derive across country genetic correlation estimates. Ignore newly estimated country specific (co)variances and update covariances between countries using the correlation estimates and original country parameters.

Repeat *Step 1* through *Step 4* until all across country genetic correlation estimates no longer change significantly.

2.5. Data for an application of the multi-trait MACE model

A number of traits from German national genetic evaluations are selected for validating the multi-trait MACE model. The traits include DYD of milk production traits and somatic cell scores in first three lactations, a number of linear type traits, direct functional longevity, and calving traits. Basically, all traits included the German total merit index can be considered in this application. These traits will be evaluated as if they were the same biological trait from different countries in international genetic evaluation. The three yield traits and somatic cell scores are evaluated in the national evaluations with a multiple lactation RRTDM, calving ease with a maternal effects model, the type traits with a multiple trait model, and longevity with a single trait survival model. Thus, both single and multiple trait models are covered in this application. A simulation study should be conducted to investigate accuracy and efficiency of the proposed multiple trait MACE model.

3. Discussion

A multi-trait MACE model was proposed for international bull comparison to account for the fact that some countries use a multiple trait model in national genetic evaluation, while the others a single trait model. This MACE model evaluates DYD of bulls or YD of cows together with corresponding EDC. An approximate REML algorithm was developed to estimate across country genetic correlations based on the multi-trait EDC method. For solving the large equation system of the multi-trait MACE model, pre-conditioned conjugate gradients method was applied together with the iteration on data technique. In order to validate this model, a number of traits from German national genetic evaluation were chosen and they are analysed as if they were the same biological traits from different countries in international evaluations. In this model, only across country genetic correlations are to be estimated, whereas genetic (co)variances from

individual countries remain constant during the parameter estimation. Error (co)variances within country do not need to be re-estimated either, because they are embedded in EDC of bulls.

The multi-trait MACE model [1] can be extended to evaluate DYD by production year of bulls (Ducrocq *et al.*, 2003) for integrating genetic trend validation into MACE evaluation. Multiple DYD and associated EDC are required for the extended multi-trait MACE model, and an additional regression effect on production years needs to be fitted.

Using MTEDC to approximate PEC of Mendelian sampling effects will result in more accurate PEC for the multi-trait MACE model based on DYD than reliability approximation for national evaluations based on original records, because the MACE model for DYD contains only one fixed effect plus additive genetic effects, and therefore there are much fewer off-diagonals in the coefficient matrix to be absorbed for approximation of PEC. In addition, this approximation leads to more accurate PEC with complete pedigree information traced by sire and dam of animal than by sire and maternal grandsire, because a pair of sire and maternal grandsire has more common progeny than a pair of sire and dam.

Due to the large dimension of the equation system, the iteration on data technique has to be used to avoid explicit setting up the equations in RAM. Should RAM usage of PCG for the model falls within available computing resources, quicker convergence can be expected from PCG than Gauss-Seidel or Gauss-Jacobi algorithms, as the iterative minimisation search strategy of PCG reduces the number of rounds of iteration to reach convergence. Because pre-conditioners of PCG should be as close as possible to the inverse of coefficient matrix of MME, EDC matrices from reliability approximation in the parameter estimation can be used as pre-conditioner. This should improve the convergence of PCG in comparison to the use of inverted diagonal blocks of coefficient matrix as pre-conditioners.

4. References

- Ducrocq, V., Boichard, D., Barbat, A. & Larroque, H. 2001. Implementation of an approximate multi-trait BLUP evaluation to combine production traits and functional traits into a total merit index. *52nd Annual Meeting of the European Association for Animal Production*, Budapest, Hungary.
- Ducrocq, V., Delaunay, I., Boichard, D. & Mattalia, S. 2003. A general approach for international genetic evaluations robust to inconsistencies of genetic trends in national evaluations. *Interbull Bulletin* 30, 101-111.
- Fikse, W.F., Klei, L., Liu, Z. & Sullivan, P.G. 2003. Procedure for validation of trends in genetic variance. *Interbull Bulletin* 31, 30-35.
- Liu, Z., Reinhardt, F., Bünger, A. & Reents, R. 2004. Derivation and calculation of approximate reliabilities and daughter yield deviations of a random regression test-day model for genetic evaluation of dairy cattle. *J. Dairy Sci.* (in press)
- Mrode, R.A. 1996. *Linear models for the prediction of animal breeding values*. CAB International, U.K.
- Schaeffer, L.R. 1994. Multiple-country comparison of dairy sires. *J. Dairy Sci.* 77, 2671-2678.
- Schaeffer, L.R. 2001. Multiple trait international bull comparison. *Livest. Prod. Sci.* 69, 145-153.
- Schaeffer, L.R. & Kennedy, B.W. 1986. Computing strategies for solving mixed model equations. *J. Dairy Sci.* 69, 575-579.
- Strandén, I. & Lidauer, M. 1999. Solving large mixed linear models using preconditioned conjugate gradient iteration. *J. Dairy Sci.* 82, 2779-2787.
- Sullivan, P. & Wilton, J.W. 2001. Multiple-trait MACE with a variable number of traits per country. *Interbull Bulletin* 27, 68-72.
- Sullivan, P. 2004. Pseudo-REML for large multiple-trait MACE applications. *Interbull Technical Committee Workshop*, Uppsala, Sweden.
- VanRaden, P.M. & Wiggans, G.R. 1991. Derivation, calculation, and use of national animal model information. *J. Dairy Sci.* 74, 2737-2746.