

Approximate restricted maximum likelihood and approximate prediction error variance of the Mendelian sampling effect

D Boichard¹, LR Schaeffer², AJ Lee³

¹ *Institut National de la Recherche Agronomique, Station de Génétique Quantitative et Appliquée, 78352 Jouy-en-Josas Cedex, France;*

² *Centre for Genetic Improvement of Livestock, University of Guelph, Ontario, N1G 2W1;*

³ *Agriculture Canada, Animal Research Centre, Ottawa, Ontario, K1A 0C6, Canada*

(Received 26 August 1991; accepted 14 May 1992)

Summary – In an Expectation-Maximization type Restricted Maximum Likelihood (REML) procedure, the estimation of a genetic (co-)variance component involves the trace of the product of the inverse of the coefficient matrix by the inverse of the relationship matrix. Computation of this trace is usually the limiting factor of this procedure. In this paper, a method is presented to approximate this trace in the case of an animal model, by using an equivalent model based on the Mendelian sampling effect and by simplifying its coefficient matrix and its inversion. This approximation appeared very accurate for low heritabilities but was downwards biased when the heritability was high. Implemented in a REML procedure, this approximation reduced dramatically the amount of computation, but provided downwards biased estimates of genetic variances. Several examples are presented to illustrate the method.

variance and covariance components / restricted maximum likelihood / Mendelian sampling effect / animal model

Résumé – Approximation du maximum de vraisemblance restreinte et de la variance d'erreur de prédiction de l'aléa de méiose. Dans certaines procédures de Maximum de Vraisemblance Restreint (REML), l'estimation des composantes de (co)variance génétique implique le calcul de la trace du produit de l'inverse de la matrice des coefficients par l'inverse de la matrice de parentés, calcul qui constitue généralement le facteur limitant de ce type de procédure. Nous présentons dans cet article une méthode visant à obtenir une valeur approchée de cette trace dans le cadre d'un modèle animal, en utilisant un modèle équivalent basé sur l'aléa de méiose, en simplifiant sa matrice des coefficients et en calculant une inverse approchée. Cette approximation est très précise lorsque l'héritabilité du caractère est faible mais elle tend à sous-estimer la trace vraie lorsque l'héritabilité est

élevée. Intégrée dans une procédure de REML, cette méthode en réduit considérablement le coût mais fournit en général des valeurs sous-estimées de variance génétique. Divers exemples sont présentés à titre d'illustration.

**composante de variance et de covariance / maximum de vraisemblance restreinte /
aléa de méiose / modèle animal**

INTRODUCTION

Restricted Maximum Likelihood (REML; Patterson and Thompson, 1971) is considered as the method of choice for estimating variance and covariance components. Applied to an animal model, REML may account at least partly for assortative matings, selection over generations and selection on a correlated trait (Meyer and Thompson, 1984; Sorensen and Kennedy, 1984). Increase in computational capacities and development of new algorithms, such as the derivative-free algorithm (Graser *et al*, 1987; Meyer, 1989a, 1991) made practical application of REML possible on medium-size data sets, particularly in analyses of selection experiments. However, there are still severe limitations with large data sets or with multiple trait models when some data are missing.

Conceptually, the Expectation-Maximization (EM) algorithm, proposed by Dempster *et al* (1977) is one of the simplest, exploiting first derivative information only. An important property of EM is that variance and covariance components estimates remain within the parameter space. It is usually slow to converge, but an acceleration (Laird *et al*, 1987) can substantially reduce the number of iterations required. However, the EM algorithm requires the inverse of the coefficient matrix for random effects. More than the repeated solution of animal model equations, calculation of this inverse is the primary limitation computationally, particularly when the coefficient matrix is large. Some attempts have already been made to approximate this inverse or at least its diagonal (Wright *et al*, 1987; Tavernier, 1990) but not under an animal model with complete relationships.

The objectives of this paper were 1) to present an approximate method for computing the trace involved in an EM-type REML algorithm for an animal model with one class of fixed effects and one class of random effects, 2) to derive an approximate variance-covariance component estimation procedure suited to large data sets and some kinds of multiple trait models, and 3) to examine the accuracy of this approximate method in applications.

METHODS

Use of an equivalent model

For simplicity, the main development is described initially with a single trait model, and its extension to the multiple trait situation will be presented in a second step. Let the model be:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

with \mathbf{Y} being the vector of observations,

β being the vector of fixed effects, assumed to include only one factor called management group,

u being the vector of n additive genetic effects, with expectation $E(u) = 0$ and variance $V(u) = A\sigma_u^2$, A being the numerator relationship matrix,

e being the vector of residual effects, with expectation $E(e) = 0$, variance $V(e) = I\sigma_e^2$ and zero covariance between u and e , and X and Z being the corresponding design matrices.

In an EM-type REML, σ_u^2 is usually estimated iteratively by (Henderson, 1984):

$$\hat{\sigma}_u^{2^{[k+1]}} = \left[\hat{u}^{[k+1]}' A^{-1} \hat{u}^{[k+1]} + \text{tr}(A^{-1} C_{22}^{[k]}) \hat{\sigma}_e^{2^{[k]}} \right] / n \tag{1}$$

with C_{22} being the $n \times n$ block of the inverse C of the coefficient matrix, pertaining to genetic effects, and $[k]$ the round of iteration. In the following part, superscript $[k]$ be will omitted.

Following Henderson (1976), if the individuals are sorted from the oldest to the youngest, the inverse of the coefficient matrix can be written as:

$$A^{-1} = L'DL$$

L is a lower triangular matrix with one on the diagonal and at most 2 non-zero terms per row, equal to -0.5 and relating a progeny to its parents. D is a diagonal matrix with general term d_{ii} , with

- $d_{ii} = 4/(2 - \phi_s - \phi_d)$ if both parents s and d of i are known,
- $d_{ii} = 4/(3 - \phi_s)$ if one parent, say s , is known,
- $d_{ii} = 1$ if both parents of i are unknown,
- ϕ_s being the inbreeding coefficient of the parent s .

Quaas (1984) proposed an equivalent model based on the Mendelian sampling effect (w), ie the deviation of the progeny breeding value from parental average.

$$Y = X\beta + ZL^{-1}w + e$$

with $w = Lu$, $E(w) = 0$ and $V(w) = D^{-1}\sigma_u^2$. Meyer (1987) showed that the use of this equivalent model may simplify the estimation of variance components. The two parts of the right-hand side in [1] can be rewritten as:

$$\hat{u}' A^{-1} \hat{u} = \hat{u}' L' D L \hat{u} = \hat{w}' D \hat{w} = \sum_{i=1}^n d_{ii} \hat{w}_i^2$$

and

$$\begin{aligned} \text{tr}(A^{-1} C_{22}) &= \text{tr}[L' D L (Z' M Z + \lambda L' D L)^{-1}] \\ &= \text{tr}[D (L^{-1}' Z' M Z L^{-1} + \lambda D)^{-1}] \\ &= \text{tr}[D K^{-1}] \end{aligned}$$

with \mathbf{M} being the matrix of fixed effects absorption, λ the variance ratio at iteration k , and \mathbf{K} the coefficient matrix of the equivalent model, after absorption of the fixed effects.

$$\begin{aligned}\mathbf{M} &= \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \\ \lambda &= \sigma_e^2/\sigma_u^2 \\ \mathbf{K} &= \mathbf{L}^{-1}'\mathbf{Z}'\mathbf{M}\mathbf{Z}\mathbf{L}^{-1} + \lambda\mathbf{D}.\end{aligned}$$

Therefore

$$\text{Var}(\mathbf{w} - \hat{\mathbf{w}}) = \mathbf{K}^{-1}\sigma_e^2$$

Because \mathbf{D} is diagonal, only the diagonals of \mathbf{K}^{-1} are needed to calculate $\text{tr}[\mathbf{D}\mathbf{K}^{-1}]$, and, noting that those are equal to the prediction error variances of the Mendelian sampling effects, [1] can be rewritten again as follows:

$$\hat{\sigma}_u^2 = \left[\sum_{i=1}^n d_{ii} (\hat{w}_i^2 + \text{var}(\hat{w}_i - w_i)) \right] / n$$

The next step is to determine the prediction error variance of the individual Mendelian sampling effects or, equivalently, the diagonal of \mathbf{K}^{-1} .

Simplification of $\mathbf{K} = \mathbf{L}^{-1}'\mathbf{Z}'\mathbf{M}\mathbf{Z}\mathbf{L}^{-1} + \lambda\mathbf{D}$

\mathbf{L}^{-1} is a lower triangular matrix with general term l_{ij} being the expected proportion of i 's genes coming from j . On the diagonal, $l_{ii} = 1$. If i is a descendant of j and n the number of generations between i and j , then $l_{ij} = \Sigma 0.5^n$; $l_{ij} = 0$ otherwise. If j appears several times in the pedigree of i , the contributions are summed over the different pathways. In absence of inbreeding, $l_{ij} = 0.5$ if i is a progeny of j , 0.25 if i is a grand progeny of j , and so on. The structure of \mathbf{K} may be examined. Its general term k_{ij} can be written as

$$k_{ij} = \left(\sum_{k=i}^n \sum_{l=j}^n l_{ki} l_{lj} z_{kl} \right) + \lambda d_{ij}$$

with d_{ij} being the general term of \mathbf{D} ($d_{ij} = 0$ if i different to j) and z_{kl} the general term of $\mathbf{Z}'\mathbf{M}\mathbf{Z}$. Accordingly, k_{ij} is non-zero if one of the 4 following conditions is fulfilled: i and j are related; or i and j are contemporary (*ie* have a record in the same management group); or i and j have a common descendant; or both i and j have a descendant, and these 2 descendants are contemporary. Consequently, the \mathbf{K} matrix is rather dense and the non-zero proportion is frequently over 50%. Therefore, its exact inverse is computationally expensive to obtain and 2 simplifications are proposed to derive a sparse approximate \mathbf{K} matrix.

The covariance between contemporaries, generated by the management group absorption is assumed to be null. Consequently, $\mathbf{Z}'\mathbf{M}\mathbf{Z}$ remains diagonal with general term z_{ii} equal to $1 - 1/n_h$, if i has a record, with n_h the number of records

in the management group h of i . Off-diagonal terms of $\mathbf{Z}'\mathbf{M}\mathbf{Z}$, equal to $-1/n_h$, are neglected. Obviously, the smaller n_h , the greater the impact of this simplification.

Only the diagonals (1) and the first-order terms relating parents to progeny (0.5) of \mathbf{L}^{-1} are taken into account, and the other terms are neglected.

After these 2 simplifications, the density of \mathbf{K} is very low and its structure is simple. That is, an individual may be related with a non-zero term in \mathbf{K} only to its parents, its progeny and its mates. Its structure looks like that of \mathbf{A}^{-1} (Henderson, 1976) and consequently \mathbf{K} may be obtained directly from a pedigree list and a data file, according to the following rules. Assuming z_{ii} equal to 0 for animals without records and $(1 - 1/n_h)$ for animals with a record, contributions to \mathbf{K} of animal i , with sire s and dam d , are the following:

<i>Add</i>	<i>To element</i>	<i>Condition</i>
$z_{ii} + \lambda d_{ii}$	k_{ii}	
$1/4 z_{ii}$	k_{ss}	if s is known
$1/2 z_{ii}$	k_{is}, k_{si}	if s is known
$1/4 z_{ii}$	k_{dd}	if d is known
$1/2 z_{ii}$	k_{id}, k_{di}	if d is known
$1/2 z_{ii}$	k_{sd}, k_{ds}	if both s and d are known

Approximate inversion of \mathbf{K}

More exactly, only the diagonal of \mathbf{K}^{-1} is needed. *A priori* the structure of the \mathbf{K} matrix is rather favourable since only the diagonal terms receive contributions of the variance ratio λ , weighted by d_{ii} , which is greater than or equal to one. Therefore, the diagonal terms are consistently higher than the off-diagonals, particularly when the variance ratio is high, *ie* when the heritability is low. Schaeffer (1990) proposed an approximation of the diagonal of the inverse by the inverse of the diagonal terms of \mathbf{K} . According to the structure of \mathbf{K} , similar to that of \mathbf{A}^{-1} , Meyer's method (1989b) can be adapted. Meyer's method is an approximate method to obtain prediction error variances of breeding values under an animal model. The basic idea is to adjust diagonal terms of each individual in the mixed model equations, by absorbing relatives equations, and to invert the resulting term. For each animal, only the most important equations, corresponding to its parents, its progeny and its management group are formally absorbed. However, processing the pedigree in the right order makes it possible to concentrate information from the whole population to a given animal. Such a process involves 2 steps. First, the sequential absorption of progeny equations into parents, from the youngest to the oldest progeny in the population, and secondly, the sequential absorption of parents equations into progeny, in the reverse order. The same algorithm can be applied to the \mathbf{K} matrix. Let i be an animal with sire s and dam d and let k_{ii}^t and k_{ii}^{t+1} denote its diagonal term in \mathbf{K} before and after adjustment respectively.

Absorption of progeny equations into parents, from the youngest to the oldest progeny, gives

$$\begin{aligned}
 k_{ss}^{t+1} &= k_{ss}^t - 1/4z_{ii}^2/k_{ii}^t && \text{if } s \text{ is known} \\
 k_{dd}^{t+1} &= k_{dd}^t - 1/4z_{ii}^2/k_{ii}^t && \text{if } d \text{ is known}
 \end{aligned}$$

Absorption of parents' equations into progeny, from the oldest to the youngest progeny, gives

$$\begin{aligned}
 k_{ii}^{t+1} &= k_{ii}^t - 1/4z_{ii}^2/k_{ss}^* && \text{if only } s \text{ is known,} \\
 k_{ii}^{t+1} &= k_{ii}^t - 1/4z_{ii}^2/k_{dd}^* && \text{if only } d \text{ is known,} \\
 k_{ii}^{t+1} &= k_{ii}^t - 1/4z_{ii}^2(k_{ss}^* + k_{dd}^* - z_{ii}/2)/(k_{ss}^*d_{dd}^* - z_{ii}^2/16)
 \end{aligned}$$

if both s and d are known, with k_{ss}^* and k_{dd}^* being the diagonal terms corresponding to parents, after disadjustment for i 's information, *ie.*

$$\begin{aligned}
 k_{ss}^* &= k_{ss}^t + 1/4z_{ii}^2/k_{ii}^t \\
 k_{dd}^* &= k_{dd}^t + 1/4z_{ii}^2/k_{ii}^t
 \end{aligned}$$

Then the i th diagonal term of \mathbf{K}^{-1} is approximated by $1/k_{ii}$.

Extension to multiple trait models

Consider now a model with q traits, possibly with missing data. Let \mathbf{G} be the non singular $q \times q$ genetic variance-covariance matrix and \mathbf{G}^{-1} its inverse. Let \mathbf{R}_i^- be a generalized inverse of the $q \times q$ residual variance-covariance matrix corresponding to individual i , with null rows and columns according to missing data. Firstly, \mathbf{R}_i^- is adjusted for the fixed effect absorption:

$$\mathbf{R}_i^{-*} = \mathbf{R}_i^- - \mathbf{R}_i^- \mathbf{W}_h^{-1} \mathbf{R}_i^-$$

and

$$\mathbf{W}_h = \sum_{i=1}^{n_h} \mathbf{R}_i^-$$

If \mathbf{K}_{ij} is the $q \times q$ block of the \mathbf{K} matrix corresponding to animals i and j , the rules to build the \mathbf{K} matrix are similar to those in part *B*. Contributions of animal i , with sire s and dam d , are the following:

Add	To block	Condition
$\mathbf{R}_i^{-*} + d_{ii}\mathbf{G}^{-1}$	\mathbf{K}_{ii}	
$1/4 \mathbf{R}_i^{-*}$	\mathbf{K}_{ss}	if s is known
$1/2 \mathbf{R}_i^{-*}$	$\mathbf{K}_{is}, \mathbf{K}_{si}$	if s is known
$1/4 \mathbf{R}_i^{-*}$	\mathbf{K}_{dd}	if d is known
$1/2 \mathbf{R}_i^{-*}$	$\mathbf{K}_{id}, \mathbf{K}_{di}$	if d is known
$1/4 \mathbf{R}_i^{-*}$	$\mathbf{K}_{sd}, \mathbf{K}_{ds}$	if both s and d are known

Again, strategies of Schaeffer and Meyer can be applied. In the first one, off-diagonals blocks \mathbf{K}_{ij} are neglected and the \mathbf{K}_{ii} blocks are inverted. With Meyer's method, the 3 steps are the following:

Absorption of progeny equations into parents, from the youngest to the oldest progeny in the population, gives

$$\begin{aligned} \mathbf{K}_{ss}^{t+1} &= \mathbf{K}_{ss}^t - 1/4 \mathbf{R}_i^{-*} \mathbf{K}_{ii}^{-1} \mathbf{R}_i^{-*} && \text{if } s \text{ is known,} \\ \mathbf{K}_{dd}^{t+1} &= \mathbf{K}_{dd}^t - 1/4 \mathbf{R}_i^{-*} \mathbf{K}_{ii}^{-1} \mathbf{R}_i^{-*} && \text{if } d \text{ is known,} \end{aligned}$$

Absorption of parents equations into progeny, in the reverse order, is performed using one of the formulae, according to whether one or both parents are known. If one parent, say s , is known,

$$\mathbf{K}_{ii}^{t+1} = \mathbf{K}_{ii}^t - 1/4 \mathbf{R}_i^{-*} \mathbf{V}_s^{-1} \mathbf{R}_i^{-*}$$

with

$$\mathbf{V}_s = \mathbf{K}_{ss}^t + 1/4 \mathbf{R}_i^{-*} \mathbf{K}_{ii}^{-1} \mathbf{R}_i^{-*}$$

If both parents are known

$$\mathbf{K}_{ii}^{t+1} = \mathbf{K}_{ii}^t - 1/16 \begin{bmatrix} \mathbf{R}_i^{-*} & \mathbf{R}_i^{-*} \end{bmatrix} \mathbf{V}^{-1} \begin{bmatrix} \mathbf{R}_i^{-*} \\ \mathbf{R}_i^{-*} \end{bmatrix}$$

with

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_s & 1/4 \mathbf{R}_i^{-*} \\ 1/4 \mathbf{R}_i^{-*} & \mathbf{V}_d \end{bmatrix}$$

Finally, invert the \mathbf{K}_{ii} blocks.

Material

The accuracy of the present method was investigated at 2 different levels. First, the approximate trace $\text{tr}(\mathbf{A}^{-1} \mathbf{C}_{22})$ was compared to the true one. Three different data sets were used. The first one was a small simulated data set with 150 animals over 5 generations and records in 17 management groups. It was used to measure the effect of each individual simplification (\mathbf{L}^{-1} , management group absorption, inversion). The other 2 data sets, of medium size, corresponded to real examples. The "cattle" data set included 722 feed efficiency records of Holstein heifers of the Agriculture Canada experimental farm in Ottawa. Records were distributed in 44 management groups and, after adding pedigree information, 1 248 animals were evaluated. The "chicken" data set included residual feed intake (\mathbf{R}) data of a chicken line, called $\mathbf{R}-$ and selected over 15 discrete generations (Bordas and Mérat, 1984). This line included 2 620 chickens and 640 parents with a complex family structure. In these 3 situations, approximate traces obtained according Schaeffer's and Meyer's strategies were compared to the true trace under 4 heritabilities (0.01, 0.10, 0.25, 0.50).

At the second level, an approximate REML was implemented and compared to a true one. Results were based on the chicken data. The female residual feed intake

(\mathbf{R}) was defined as the deviation of observed feed intake from a theoretical feed intake predicted from maintenance, change in body weight and egg production. For the male trait, only maintenance and change in body weight were accounted for. Firstly, the female residual feed intake was analyzed alone in a single trait animal model. Next, because preliminary results led us to assume that the male and the female \mathbf{R} were not the same trait, they were analysed in a 2 trait model. To decrease the computation cost of the true REML, and particularly the bivariate one, requiring repeated inversion of the reduced animal model coefficient matrix, the first 12 generations only were analysed. The characteristics of the data set are in table I. To speed up convergence, an exponential acceleration (Laird *et al*, 1987) was used every 6 iterations but was applied only if the resulting variance-covariance matrices were positive definite.

Table I. Characteristics of the “chicken” data set used in the REML analysis.

	<i>Single trait analysis</i>	<i>Multitrait analysis</i>
Animals evaluated	1 236	1 486
Generations	12	12
Female records	1 153	1 153
Male records	–	314
Female fixed effects	20	20
Male fixed effects	–	17
Parents	376	383

RESULTS

Comparison of true and approximate traces

Table II shows the results obtained from the small simulated data set. The density of \mathbf{K} was strongly reduced from 39.4% without approximation to 2.9% with simplifications of \mathbf{L}^{-1} and management group absorption. This reduction is expected to be much more important in large applications since the number of non-zero terms in the approximate coefficient matrix \mathbf{K} is less than 7 times the number of animals.

Obviously, the true trace increased with heritability, because the prediction error variance of each Mendelian sampling effect increases with genetic variability. Generally, the simplification of \mathbf{L}^{-1} led to a small increase of the trace, while the simplification of the management group absorption led to a decrease, particularly for high values of heritability. This example was rather unfavourable to the simplified methods since the average number of contemporaries n_h was rather small (8), and moreover, contemporaries were often highly related.

The approximate inversion of \mathbf{K} had no additional effect when the heritability was low but led to underestimating the trace when the heritability was high, and this bias was larger with Schaeffer’s method, *ie* when off-diagonal terms were neglected, than with Meyer’s. When the heritability is low, the variance ratio λ is high and

Table II. Effect of different approximations on $\text{tr}(\mathbf{A}^{-1}\mathbf{C}_{22})$ in the simulated data set.

Methods	Simplification		Inverse	Density of \mathbf{K} (%)	Heritability			
	\mathbf{L}^{-1}	Absorption			0.01	0.10	0.25	0.50
(1) = true	No	No	Inversion	39.4	^b 1.504 4	15.552	42.06	103.22
(2)	Yes	No	Inversion	24.8	1.505 2	15.600	42.18	102.89
(3)	No	Yes	Inversion	7.4	1.503 8	15.495	41.65	100.15
(4)	Yes	Yes	Inversion	2.6	1.504 8	15.557	41.80	99.75
(5)	Yes	Yes	Schaeffer ^a	2.6	^b 1.504 8	15.557	41.17	93.03
					^c + 0.023%	-0.21%	-2.12%	-9.87%
(6)	Yes	Yes	Meyer	2.6	1.504 8	15.552	41.73	99.21
					+0.025%	-0.003%	-0.78%	-3.88%

^a $\text{Diag}(\mathbf{K}^{-1}) = [\text{Diag}(\mathbf{K})]^{-1}$; ^b value of $\text{tr}(\mathbf{A}^{-1}\mathbf{C}_{22})$; ^c bias in percent of true value of $\text{tr}(\mathbf{A}^{-1}\mathbf{C}_{22})$.

\mathbf{K} 's off-diagonal terms are much lower than the diagonals and can be neglected. With a high heritability, this is no longer the case and Schaeffer's methods becomes clearly less efficient than Meyer's method. Finally, when the 3 approximations were accumulated and when the heritability was low, $\text{tr}(\mathbf{A}^{-1}\mathbf{C}_{22})$ was well approximated by both methods, generally differing by much less than 1% from true value. When heritability increased, Meyer's method appeared more efficient than Schaeffer's but still underestimated $\text{tr}(\mathbf{A}^{-1}\mathbf{C}_{22})$.

Results for the larger data sets ("chicken" in table III and "cattle" in table IV) were basically the same. In the "cattle" data set with Meyer's method, the bias was slightly positive (0.09 to 0.55%) for a low or medium heritability and slightly negative (-0.51%) for a high heritability. This good result is probably related to the small number of generations and the large average number of contemporaries.

Table III. Comparison of 3 methods to compute $\text{tr}(\mathbf{A}^{-1}\mathbf{C}_{22})$ in the "cattle" data set.

Heritability	True method ¹	Schaeffer ² (%)	Meyer ² (%)
0.01	12.5413	+0.09	+0.09
0.10	132.13	+0.39	+0.54
0.25	369.73	-0.38	+0.55
0.50	972.60	-5.02	-0.51

¹ Value of $\text{tr}(\mathbf{A}^{-1}\mathbf{C}_{22})$. ² Bias in percent of the true value.

In the "chicken" data set, bias was generally negative and reached -2.19% when heritability was 0.05. This result, less favourable than in the previous example, is probably due to the number of generations and to the relatively small number of reproducers. In spite of a large average number of contemporaries, the effect of the

Table IV. Comparison of 3 methods to compute $\text{tr}(\mathbf{A}^{-1}\mathbf{C}_{22})$ in the “chicken” data set.

<i>Heritability</i>	<i>True method</i> ¹	<i>Schaeffer</i> ² (%)	<i>Meyer</i> ² (%)
0.01	26.24	+0.09	+0.09
0.10	270.26	+0.23	-0.02
0.25	727.12	-2.02	-0.84
0.50	1 719.96	-6.85	-2.19

¹ Value of $\text{tr}(\mathbf{A}^{-1}\mathbf{C}_{22})$. ² Bias in percent of the true value.

absorption simplification was inflated because contemporaries were related, at least after several generations (the average inbreeding coefficient at the last generation was 0.28).

In both data sets with Schaeffer’s method, the bias was very small for a low heritability but reached -5.02 and -6.85% with a heritability of 0.5. Therefore, in spite of its (relative) complexity, particularly in the multiple trait situation, Meyer’s method was chosen for the approximate REML analysis presented in the following part B.

REML analysis

While the computation of $\text{tr}(\mathbf{A}^{-1}\mathbf{C}_{22})$ is usually the limiting factor of the EM-type REML, its cost is negligible in the approximate REML compared to the repeated solution of animal model equations.

Table V presents the results of the female “chicken” data analysis at the first iteration and at convergence. The starting value for the variance ratio was the same (3) in the true REML analysis and in the approximate one. At the first iteration, the contribution of the prediction error variances $\text{tr}(\mathbf{A}^{-1}\mathbf{C}_{22})$ appeared 6 times larger than the contribution of the quadratic form of the estimated breeding values. Under this very unfavourable situation and with the approximate method, the bias in the estimation of the trace was almost undiluted and led to an almost equivalent bias in the estimate of the variance component. The bias in the trace estimation was rather small at any one given iteration, for example -0.64% at the first and -0.40% at the convergence point of the true REML. However, the bias was accumulated over iterations and the heritability estimate at convergence was clearly underestimated (0.173 *vs* 0.208). These estimates were independent of the starting value.

Results of the bivariate analysis of the “chicken” data are presented in table VI. They were basically the same as for the single trait analysis. At convergence, the estimates of the approximate method were found to be always the same, regardless of starting values. The trace $\text{tr}(\mathbf{A}^{-1}\mathbf{C}_{22})$ was underestimated, particularly for the male trait, which was the most heritable and with the smallest average number of contemporaries n_h (18.5 *vs* 57.6 for the female trait). At convergence of the true REML, the absolute approximate trace was underestimated by -0.53% for the male trait (with heritability 0.57), by -0.33% for the female trait (with heritability 0.21) and by -0.29% for the combination of both traits, with an almost zero

Table V. Results from the single trait REML analysis (female "chicken" data).

	<i>True REML</i>	<i>REML</i>	
First iteration ¹			
$\mathbf{u}' \mathbf{A}^{-1} \mathbf{u}$	2 082 460	2 082 460	
$\text{tr}(\mathbf{A}^{-1} \mathbf{C}) \sigma_e^2$	11 295 373	11 223 477	(-0.64%)
At convergence			
σ_u^2	8 833	7 303	(-17.3%)
σ_e^2	33 816	34 819	(+2.97%)
h^2	0.207	0.173	(-20.2%)
$\text{tr}(\mathbf{A}^{-1} \mathbf{C}) \sigma_e^2$ for $h^2 = 0.207$	9 480 696	9 443 121	(-0.40%)

¹ Value of $\sigma_e^2 / \sigma_u^2 = 3$.

Table VI. Results from the bivariate REML analysis (male and female "chicken" data).

	<i>True REML</i>	<i>Approximate REML</i>	
Genetic parameters			
σ_{um}^2	58 233	38 799	(-33.7%)
σ_{uf}^2	8 858	7 309	(-17.4%)
σ_{umf}	-808	-505	(+37.5%)
σ_{em}^2	42 337	54 158	(+27.9%)
σ_{ef}^2	33 788	34 804	(+3.0%)
h_m^2	0.579	0.417	
h_f^2	0.208	0.174	
r_g	-0.036	-0.030	
Traces at convergence of the true REML			
male trait	78 870 651	78 453 231	(-0.53%)
female trait	11 719 000	11 680 853	(-0.33%)
covariance	-966 278	-963 500	(+0.29%)

genetic correlation (-0.04). Consequently, a clearly different convergence point was reached with the approximate method. The genetic variance components were underestimated, resulting in a strong downwardly biased estimate of the male heritability (0.417 vs 0.579), a moderately biased estimate of female heritability (0.174 vs 0.208) and an almost unbiased estimate (probably by chance) of genetic correlation (-0.03 vs -0.04).

DISCUSSION AND CONCLUSION

Although the approximate method gives rather accurate estimates of the prediction error variance contribution at any one iteration, it does not provide satisfactory results in the REML analysis. This apparent contradiction is explained by the properties of the animal model. The variance components are estimated as the sum of the quadratic form of the breeding values, which is the really informative part, and the prediction error variance, which should be only an adjustment factor. In the case of an animal model, the amount of information carried by each animal is much smaller than the adjustment factor. In this unfavourable situation, a small bias in estimating this adjustment factor estimate leads to a variance estimate which may not be very close to the REML solution. Because the accuracy of the Mendelian sampling effect estimate is not primarily dependent of the population size, this problem is not expected to be solved by increasing the size of the data sample. To a lesser extent, similar problems may arise in a true REML when $\text{tr}(\mathbf{A}^{-1}\mathbf{C}_{22})$ is not computed accurately enough, because of rounding errors in the inversion of large coefficient matrices. This may explain differences in results between methods or algorithms, or some surprising convergence points (Groeneveld and Kovac, 1990).

To avoid this problem, 2 ways might be investigated in further analysis. One approach would be to develop a similar method suited to models with fewer animals involved, each concentrating more information, as for instance in the reduced animal model (Quaas and Pollack, 1980). In that case, the ratio $\mathbf{u}'\mathbf{A}^{-1}\mathbf{u}/\text{tr}(\mathbf{A}^{-1}\mathbf{C}_{22})$ would be increased and the method would be more robust to any bias in $\text{tr}(\mathbf{A}^{-1}\mathbf{C}_{22})$. Another way would be to quantify by simulation the effect on the bias of heritability, the distribution of the data in the contemporary groups and the family structure, in order to adjust the trace *a priori*.

Presently, the approximate method does not provide the same estimates as a true REML, and further developments are needed to make it more efficient. Although the examples presented here are not general, it can be concluded that the bias increases when heritability increases, when the size of contemporary group decreases and when animals in the same contemporary group are related. Owing to its ease of use, this approximate method can be recommended only as a first approach, when the true heritability is expected to be low or moderate and when the contemporary groups are large. However, its use is restricted to the class of models in which the residual components can be computed as a residual sum of squares. Until now, no approximation has been found for the residual components in the general case of multiple trait models with missing data.

ACKNOWLEDGMENT

This work was carried out when the first author was at the Animal Research Centre (Ottawa), with the financial support of INRA.

REFERENCES

- Bordas A, Mérat P (1984) Correlated responses in a selection experiment of residual feed intake of adult Rhode Island Red cocks and hens. *Ann Agric Fenn* 23, 233-237
- Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data with the EM algorithm. *J R Stat Soc B* 39, 1-22
- Graser HU, Smith SP, Tier B (1987) A derivative-free approach for estimating variance components in animal models by restricted maximum likelihood. *J Anim Sci* 64, 1362-1370
- Groeneveld E, Kovac M (1990) A note on multiple solutions in multivariate maximum likelihood covariance components estimation. *J Dairy Sci* 73, 2221-2229
- Henderson CR (1976) A simple method to compute the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* 32, 69-83
- Henderson CR (1984) Estimation of variances and covariances under multiple trait models. *J Dairy Sci* 67, 1581-1589
- Laird N, Lange N, Stram D (1987) Maximum likelihood computations with repeated measures: application of the EM algorithm. *J Am Stat Assoc* 82, 97-105
- Meyer K (1987) A note on the use of an equivalent model to account for relationships between animals in estimating variance components. *J Anim Breed Genet* 104, 163-168
- Meyer K (1989a) Restricted maximum likelihood to estimate variance components for animal models with several random effects using a derivative-free algorithm. *Genet Sel Evol* 21, 317-340
- Meyer K (1989b) Approximate accuracy of genetic evaluation under an animal model. *Livest Prod Sci* 21, 87-100
- Meyer K (1991) Estimating variances and covariances for multivariate animal models by restricted maximum likelihood. *Genet Sel Evol* 23, 67-83
- Meyer K, Thompson R (1984) Bias in variance and covariance estimators due to selection on a correlated trait. *Z Tierz Zuechtungsbiol* 101, 33-50
- Patterson HD, Thompson R (1971) Recovery of inter-block information when block sizes are unequal. *Biometrika* 58, 545-554
- Quaas RL (1984) Linear prediction. In: *BLUP School Handbook*. Univ New England, Armidale, Australia
- Quaas RL, Pollack EJ (1980) Mixed model methodology for farm and ranch beef cattle testing programs. *J Anim Sci* 51, 1277-1287
- Schaeffer LR (1990) Problems in estimating covariance matrices in multitrait animal model. In: *Proc 4th World Congress of Genetics Applied to Livestock Production, Edinburgh, July 23-27 1990* (Hill WG, Thompson R, Woolliams JA, eds) Edinburgh, UK, vol 13, 441-444
- Sorensen DA, Kennedy BW (1984) Estimation of genetic variances from unselected and selected populations. *J Anim Sci* 59, 1213-1223
- Tavernier A (1990) Estimation of breeding value of jumping horses from their rank. *Livest Prod Sci* 26, 277-290
- Wright HB, Pollack EJ, Quaas RL (1987) Estimation of variance and covariance components to determine heritabilities and repeatability of weaning weight in American Simmental cattle. *J Anim Sci* 65, 975-981