On parsimonious and equivalent animal models with (grand) maternal effects and missing (grand) dams ☆

M.J. Suárez [a], A.N. Birchmeier [a], R.J.C. Cantet [a,b,*]

[a] Departamento de Producción Animal, Universidad de Buenos Aires (UBA), Avenida San Martín 4453, C1417DSE Buenos Aires, Argentina
[b] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

ARTICLE INFO

ABSTRACT

In breeds where a large fraction of animals with records on a maternally affected trait are from dams that have no records and unknown parents, the genetic evaluation of such trait may be hindered by misspecification of the genetic covariance matrix. The specified covariance structure for the additive direct and maternal effects in the regular maternal animal model (MAM) when dams have no records differs from the covariance between relatives with maternal effects. Two solutions are possible. One is to include in the vectors of breeding values for direct and maternal effects the dam or a "phantom" dam if the latter is unknown. As a consequence, the number of equations to be solved in the MAM may increase considerably. Alternatively, one may replace the maternal breeding value of the dam with 2/3 of the maternal breeding of the individual, and −1/3 of the maternal breeding value of the sire of the individual. As this "regression" of breeding values has been largely ignored, the goal of this paper is to present a parsimonious equivalent MAM using such regression. The approach is extended to a similar situation for models with grand maternal effects. Two small numerical examples are used to illustrate the proposed methods.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Newly composite beef breeds usually have an open policy of registering animals. In these breeds a large fraction of animals with records on a maternally affected trait such as weaning weight, are calved by dams that themselves have no records. Moreover, most of these dams lack pedigree information (i.e. their sire and dam identifications are missing) and usually have only one calf with records in the data base. A consequence of having dams with missing records on the genetic evaluation of a maternally affected trait is a possible misspecification of the genetic covariance matrix (Cantet et al., 1992). When an individual with record has a dam without a record, the specified covariance structure for the additive direct and maternal effects in the regular maternal animal model (MAM) is different from the covariance between relatives with maternal effects as presented by Willham (1963). Cantet et al. (1992) observed that the additive covariance between the breeding values of dam (D) and offspring (O), when the dam has no record, is equal to $1/2\sigma_{Ao}^2 + \sigma_{AoAm}$. Instead, in the formulation of Willham (1963), the covariance is equal to $1/2\sigma_{Ao}^2 + 5/4\sigma_{AoAm} + 1/2\sigma_{Am}^2$, where σ_{Ao}^2 and σ_{Am}^2 are the

additive variance for direct and maternal effects, respectively, and σ_{AoAm} is the covariance between both types of breeding values. This problem is generally ignored.

There are two possible solutions for specifying correctly the covariance between breeding values (Willham, 1963) in a MAM when some dams have no records. The simplest one is to include in the vectors of breeding values for direct and maternal effects the dam or a "phantom" dam if the latter is unknown (Westell et al., 1988; Van Vleck, 1990; Cantet et al., 1992). Clearly, the number of equations to be solved in the MAM is greater than the case where those dams are ignored. Although in some data bases this increase is inconsequential, for situations of composite breeding with a large fraction of natural matings, the number of equations can increase up to 30% or more.

While reviewing the paper of Cantet et al. (1992), Richard L. Quaas in 1991 proposed another solution. He suggested replacing the maternal breeding value of the dam (a_{mD}) with 2/3 of the maternal breeding of the individual (a_{mO}), and $-1/3$ of the maternal breeding of the sire (S) of the individual (a_{mS}). Although the method has been used ever since the genetic evaluation of beef cattle at Universidad de Buenos Aires, this "regression" of breeding values has been largely ignored. Different equivalent (Henderson, 1985) MAMs are presented when dams of recorded animals have no records on their own. The approach is extended to a similar situation in models with grand maternal effects (Willham, 1972; Dodenhoff et al., 1998, 1999a,b).

2. An equivalent maternal animal model to deal with animals with records and missing dams

The model equation for the MAM is equal to

$$y_{ijk} = \mathbf{x}_i'\boldsymbol{\beta} + a_{oi} + a_{mj} + e_{mj} + e_{oi} \tag{1}$$

In Eq. (1), y_{ijk} is the record of animal i with dam j and sire k; $\mathbf{x}_i'$ is the row vector in the incidence matrix for the vector of fixed effects ($\boldsymbol{\beta}$) corresponding to the record of animal i, a_{oi} is the direct breeding value of i, a_{mj} and e_{mj} are respectively the maternal breeding value and environmental effect of dam j, and e_{oi} is the error term. When j is unknown but sire k is known, an alternative animal model equivalent (Henderson, 1985) to (1) results from "regressing" the maternal breeding value of the dam on a linear combination of the maternal breeding values of the progeny and the sire such that

$$y_{ijk} = x_i'\boldsymbol{\beta} + a_{oi} + b_i a_{mi} + b_k a_{mk} + r_i + e_{mj} + e_{oi} \tag{2}$$

where $r_i = a_{mj} - b_i a_{mi} - b_k a_{mk}$ is a maternal genetic residual term for animal i, and b_k and b_i are the regression coefficients of a_{mj} on a_{mk} and a_{mi}, respectively. Model (2) is written more compactly as follows

$$y_{ijk} = x_i'\boldsymbol{\beta} + a_{oi} + b_i a_{mi} + b_k a_{mk} + \varepsilon_{oi} \tag{3}$$

with

$$\varepsilon_{oi} = r_i + e_{mj} + e_{oi} = a_{mj} - b_i a_{mi} - b_k a_{mk} + e_{mj} + e_{oi} \tag{4}$$

In the next section we prove that in the equivalent model (3), $b_k = -1/3$ and $b_i = 2/3$.

3. Regression of the breeding value of a dam on the breeding values of the sire and their common offspring

The solution of R. L. Quaas can be viewed as a regression of the BLUP (*Best Linear Unbiased Predictor*, Henderson, 1984) of the breeding value of the unknown dam j on the BLUPs of breeding value of progeny i and sire k. In doing so, we form the linear combination:

$$\text{BLUP}(a_j) = b_k \text{BLUP}(a_k) + b_i \text{BLUP}(a_i) \tag{5}$$

The scalars b_k and b_i are the regression coefficients of $\text{BLUP}(a_k)$ and $\text{BLUP}(a_i)$, respectively, on $\text{BLUP}(a_j)$. Although the derivation is general, when k is not inbred $b_k = -1/3$ and $b_i = 2/3$. Now, BLUP is the estimator of the conditional expectation of the random variable u given the random variable x (Henderson, 1984). Thus, we have that

$$BLUP(u) = \hat{E}(u|x) = \text{cov}(u,x)[\text{var}(x)]^{-1} BLUP(x)$$

we employ this formula for any type of breeding value (direct, maternal, grand maternal, social interaction effects, etc). Thus, $u = a_j$ and $x = [a_k, a_i]'$. Thus:

$$BLUP(a_j) = \text{cov}\left[a_j, \begin{bmatrix} a_k \\ a_i \end{bmatrix}\right] \left[Var \begin{bmatrix} a_k \\ a_i \end{bmatrix}\right]^{-1} \begin{bmatrix} \hat{a}_k \\ \hat{a}_i \end{bmatrix} = [b_k b_j] \begin{bmatrix} \hat{a}_k \\ \hat{a}_i \end{bmatrix} \tag{6}$$

The symbol ^ indicates the BLUP of the random variable. Now, by the theory of the covariance between relatives in the additive model we can write

$$\text{cov}\left[a_j, \begin{bmatrix} a_k \\ a_i \end{bmatrix}\right] = \begin{bmatrix} 0 \\ \frac{1}{2}\sigma_A^2 \end{bmatrix} \tag{7}$$

The parameter σ_A^2 is the additive variance or the variance of the breeding values. As j is an unknown dam, it is reasonable to assume that she is unrelated to the sire so that cov $(a_j, a_k)=0$. Moreover

$$Var\begin{bmatrix} a_k \\ a_i \end{bmatrix} = \begin{bmatrix} 1+F_k & \frac{1}{2}(1+F_k) \\ \frac{1}{2}(1+F_k) & 1 \end{bmatrix}\sigma_A^2$$

The inverse of this matrix is

$$\left[Var\begin{bmatrix} a_k \\ a_i \end{bmatrix}\right]^{-1} = \frac{1}{\sigma_{Ao}^2}\begin{bmatrix} \left[(1+F_k)-\frac{1}{4}(1+F_k)^2\right]^{-1} & -\left[2-\frac{(1+F_k)}{2}\right]^{-1} \\ -\left[2-\frac{(1+F_k)}{2}\right]^{-1} & \frac{1}{[1-1/4(1+F_k)]} \end{bmatrix} \tag{8}$$

After multiplying Eqs. (7) to (8), we have:

$$\begin{bmatrix} b_k \\ b_i \end{bmatrix} = \text{cov}\left[a_j, \begin{bmatrix} a_k \\ a_i \end{bmatrix}\right]Var\begin{bmatrix} a_k \\ a_i \end{bmatrix}^{-1} = \begin{bmatrix} \frac{-1}{4[1-1/4(1+F_k)]} \\ \frac{1}{2[1-1/4(1+F_k)]} \end{bmatrix} = \begin{bmatrix} \frac{-1}{[3-F_k]} \\ \frac{2}{[3-F_k]} \end{bmatrix} \tag{9}$$

when the father k is not inbred $F_k=0$; the solution corresponding to Eq. (9) is equal to:

$$\begin{bmatrix} b_k \\ b_i \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$$

When applied to Eq. (5), the resulting regression is then

$$\hat{a}_j = \frac{2}{3}\hat{a}_i - \frac{1}{3}\hat{a}_k \tag{10}$$

Regression Eq. (10) allows specifying correctly all additive (co)variances related to the records from individuals with unidentified dams in animal models.

4. Equivalence between models (1) and (3)

To prove that both models are equivalent necessitates showing that the expected values and the variances are equal (Henderson, 1985). Notice that taking expectations in both (1) and (3) produces $E\left(y_{ijk}\right) = \mathbf{x}_i'\boldsymbol{\beta}$. With regard to the variance of (3) when $b_k = -1/3$ and $b_i=2/3$, we have:

$$Var\left(y_{ijk}\right) = Var\left(a_{o\,i} + \tfrac{2}{3}a_{mi} - \tfrac{1}{3}a_{mk} + \varepsilon_{oi}\right)$$

Consider first the covariance between the genetic and residual effects, which is equal to

$$\begin{aligned}
\text{cov}&\left(a_{o\,i} + \tfrac{2}{3}a_{mi} - \tfrac{1}{3}a_{mk}, \varepsilon_{oi}\right) = \text{cov}\left(a_{o\,i} + \tfrac{2}{3}a_{mi} - \tfrac{1}{3}a_{mk}, a_{mj} - \tfrac{2}{3}a_{mi} + \tfrac{1}{3}a_{mk} + e_{mj} + e_{oi}\right) \\
&= \text{cov}\left(a_{o\,i} + \tfrac{2}{3}a_{mi} - \tfrac{1}{3}a_{mk}, a_{mj} - \tfrac{2}{3}a_{mi} + \tfrac{1}{3}a_{mk}\right) = \text{cov}\left(a_{o\,i}, a_{mj}\right) - \tfrac{2}{3}\text{cov}\left(a_{o\,i}, a_{mi}\right) \\
&\quad + \tfrac{1}{3}\text{cov}(a_{o\,i}, a_{mk}) + \tfrac{2}{3}\text{cov}\left(a_{m\,i}, a_{mj}\right) - \tfrac{4}{9}\text{cov}(a_{m\,i}, a_{mi}) + \tfrac{2}{9}\text{cov}(a_{m\,i}, a_{mk}) \\
&\quad - \tfrac{1}{3}\text{cov}\left(a_{m\,k}, a_{mj}\right) + \tfrac{2}{9}\text{cov}\left(a_{m\,k}, a_{mi}\right) - \tfrac{1}{9}\text{cov}(a_{m\,k}, a_{mk}) = \left[(1)\tfrac{1}{2} - \tfrac{2}{3}(1) + \tfrac{1}{3}\left(\tfrac{1}{2}\right)\right]\sigma_{AoAm} \\
&\quad + \left(\tfrac{2}{3}\left(\tfrac{1}{2}\right) - \tfrac{4}{9}(1) + \tfrac{2}{9}\left(\tfrac{1}{2}\right) + \tfrac{1}{3}(0) + \tfrac{2}{9}\left(\tfrac{1}{2}\right) - \tfrac{1}{9}(1)\right)\sigma_{Am}^2 = 0
\end{aligned}$$

Then, on using this result we have that

$$\begin{aligned}
Var(\varepsilon_{o\,i}) &= Var\left(a_{mj} - \tfrac{2}{3}a_{mi} + \tfrac{1}{3}a_{mk} + e_{mj} + e_{oi}\right) = Var\left(a_{mj} - \tfrac{2}{3}a_{mi} + \tfrac{1}{3}a_{mk}\right) + Var(e_{mj}) + Var(e_{oi}) \\
&= \left(1 + \tfrac{4}{9} + \tfrac{1}{9} - \tfrac{2}{3} - \tfrac{2}{9}\right)\sigma_{Am}^2 + \sigma_{Em}^2 + \sigma_{Eo}^2 = \tfrac{2}{3}\sigma_{Am}^2 + \sigma_{Em}^2 + \sigma_{Eo}^2
\end{aligned} \tag{11}$$

moreover

$$\begin{aligned}
\text{var}\left(y_{ijk}\right) &= \text{var}(a_{0i}) + \tfrac{4}{9}\text{var}(a_{mi}) + \tfrac{1}{9}\text{var}(a_{mk}) + \tfrac{4}{3}\text{cov}(a_{0i}, a_{mi}) - \tfrac{2}{3}\text{cov}(a_{0i}, a_{mk}) - \tfrac{4}{9}\text{cov}(a_{mi}, a_{mk}) \\
&\quad + \text{var}(\varepsilon_{0i}) = \sigma_{Ao}^2 + \tfrac{4}{9}\sigma_{Am}^2 + \tfrac{1}{9}\sigma_{Am}^2 + \left(\tfrac{4}{3}\right)\sigma_{AoAm} + 2\left(-\tfrac{1}{3}\right)\tfrac{1}{2}\sigma_{AoAm} + 2\left(\tfrac{2}{3}\right)\left(-\tfrac{1}{3}\right)\tfrac{1}{2}\sigma_{Am}^2 + Var(\varepsilon_{oi}) \\
&= \sigma_{Ao}^2 + \sigma_{AoAm} + \tfrac{1}{3}\sigma_{Am}^2 + Var(\varepsilon_{oi})
\end{aligned} \tag{12}$$

Inspection of Eq. (11) suggests that the quantity $\tfrac{2}{3}\sigma_{Am}^2 + \sigma_{Em}^2$ should be added to the error variance in the residual term. This will allow the variance of model (3) (i.e. Eq. (12)) to account for all genetic and permanent environmental variation

that is present in model (1), and the record of an individual with unknown dam will account for the same phenotypic variance as in animals with records and known dams.

5. An equivalent maternal animal model for individuals with both parents unknown

In case both parents are missing a different MAM is possible by replacing the maternal breeding value of dam j with half the maternal breeding value of its ith offspring:

$$a_{mj} = \frac{1}{2} a_{mi} \tag{13}$$

Thus, the resulting MAM is equal to:

$$y_{ijk} = x_i'\boldsymbol{\beta} + a_{o\ i} + \frac{1}{2} a_{mi} + \varepsilon_{oi} \tag{14}$$

Using the variance operator in Eq. (14) we obtain

$$
\begin{aligned}
Var\left(y_{ijk}\right) &= Var(a_{oi}) + Var(\tfrac{1}{2}a_{mi}) + Var(\varepsilon_{oi}) + 2\ \mathrm{cov}(a_{oi}, \tfrac{1}{2}a_{mi}) \\
&= \sigma_{Ao}^2 + \tfrac{1}{4}\sigma_{Am}^2 + \sigma_{Eo}^2 + 2\ (\tfrac{1}{2})\sigma_{AoAm} = \sigma_{Ao}^2 + \tfrac{1}{4}\sigma_{Am}^2 + \sigma_{Eo}^2 + \sigma_{AoAm}
\end{aligned} \tag{15}
$$

Thus, the quantity $\frac{3}{4}\sigma_{Am}^2 + \sigma_{Em}^2$ should be added to the error variance and Eq. (14) is a MAM with a covariance structure consistent with the formulation of Willham (1963). Therefore, the error variance is

$$Var(\varepsilon_{o\ i}) = \sigma_{Eo}^2 + \frac{3}{4}\sigma_{Am}^2 + \sigma_{Em}^2 \tag{16}$$

The regression of a_{mj} on $a_{m\ i}$ uses the regression coefficient:

$$b = \frac{\mathrm{cov}(a_{mi}, a_{mj})}{Var(a_{mi})} = \frac{(\sigma_{Am}^2/2)}{\sigma_{Am}^2} = \frac{1}{2}$$

Clearly, half the maternal breeding value of an individual contains information on the breeding value of its dam. A small numerical example of equivalent models with maternal effects and missing dams, or with both parents unknown, is presented in Appendix A.

6. An equivalent grand maternal animal model for individuals with unknown grand dams

Consider now a model including direct, maternal and grand maternal effects (GMAM, Willham, 1972; Dodenhoff et al., 1998, 1999a,b). We look at the case when the dam of an animal with record is known but the maternal grand dam is unknown. Grand maternal effects are denoted with the subscript "n". In the GMAM, the expression that is similar to Eq. (10) is

$$a_{ng} = \frac{2}{3} a_{nj} - \frac{1}{3} a_{nl} \tag{17}$$

In Eq. (17), l is the maternal grandsire of i. Whereas an expression similar to Eq. (13) is equal to

$$a_{ng} = \frac{1}{2} a_{nj} \tag{18}$$

Although expressions (17) and (18) allow for a correct specification of all additive covariances in the records of individuals with missing maternal grand dams, there exists a fraction of additive variance that is unaccounted by the model and should be added to the error term. As a result, when the breeding value of the unknown maternal grand dam g in the record of her grandson is replaced by the breeding values of her daughter and the sire of this latter dam (l), the term $\frac{2}{3}\sigma_{An}^2$ (the variance of grand maternal breeding values) should be added to the variance of the error term (ε_{oi}) such that:

$$Var(\varepsilon_{oi}) = \sigma_{Eo}^2 + \frac{2}{3}\sigma_{An}^2 \tag{19}$$

Alternatively, when the breeding value of the unknown maternal grand dam g in the record of her grandson is replaced by the breeding values of her daughter j, $\frac{1}{2}a_{nj}$, it is necessary to add $\frac{3}{4}\sigma_{An}^2$ to $Var(\varepsilon_{oi})$ for the variance of the record to be correctly specified. Then

$$Var(\varepsilon_{o\ i}) = \sigma_{Eo}^2 + \frac{3}{4}\sigma_{An}^2 \tag{20}$$

Detailed derivations of Eqs. (19) and (20) are displayed in Appendix B.

Expressions (19) and (20) are valid as long as the maternal grand sire l is not inbred. If this is not so and on denoting with F_l the inbreeding of l, the fraction of additive variance in the error term will be smaller than in Eq. (19) and

proportional to

$$\frac{1}{3}(2-F_l)\sigma_{An}^2 \tag{21}$$

A similar expression to (20) is obtained when sire k is inbred

$$\frac{1}{4}(3-F_k)\sigma_{An}^2 \tag{22}$$

After enlarging Eq. (1) to accommodate grand maternal additive effects, and when the grand maternal breeding value for unknown maternal grand dam g is replaced by a linear combination of the grand maternal breeding values of her daughter j and the maternal grand sire l, the resulting GMAM is

$$y_{ijl} = x_i'\boldsymbol{\beta} + a_{oi} + a_{mj} + \frac{2}{3}a_{nj} - \frac{1}{3}a_{nl} + e_{mj} + e_{oi} \tag{23}$$

Using a similar reasoning, when both parents of the dam of the animal with record are unknown, an equivalent GMAM with a specification of the additive covariance–matrix consistent with the formulation of Willham (1972) is equal to

$$y_{ij} = x_i'\boldsymbol{\beta} + a_{oi} + a_{mj} + \frac{1}{2}a_{nj} + e_{mj} + \varepsilon_{oi} \tag{24}$$

A small numerical example of the equivalent models with grand maternal effects and missing grand dams is presented in Appendix C.

7. Discussion

To specify correctly the additive (co)variances in the records of animals with unknown dams, Van Vleck (1990) suggested including the breeding values of "phantom" dams into the general vector of breeding values, by assuming that these females are unrelated to each other and by considering that these dams have only one grand progeny each. However, for populations evaluated with a large proportion of missing dams (i.e. newly formed breeds and composite breeds), the number of mixed model equations that have to be solved is sizable. Each phantom dam increases the number of equations by 3 in both the MAM and the GMAM. Alternatively, we present an alternative approach after the idea of R.L. Quaas that does not require increasing the number of equations. The idea is to express the breeding values of the unknown dam, as a linear function (i.e. a "regression") of the breeding values of the progeny (i.e. Eq. (10) or Eq. (13)) with record and the sire of the progeny (i.e. Eq. (10)). In the same way but for the GMAM, the breeding value of an unknown grand dam is "regressed" on the breeding values of its daughter (the dam of the animal with record, i.e. Eqs. (17) or (18)) and of the maternal grandsire of the animal with record (i.e. expression (17)). If the maternal grandsire is also unknown, then the grand maternal effect of the grand dam can be replaced by half the grand maternal effect of its daughter (see Eq. (18)). Notice that this is possible as for any individual and conditionally on the breeding values of the parents, the mendelian additive residual is independent of the breeding values of any other animal (Bulmer, 1985), and the mendelian variance and the error variance are diagonal. This facilitates an efficient way of building and solving the mixed model equations (Schaeffer and Henderson, 1983). For the models proposed here to be equivalent to those consistent with the additive covariance structures suggested by Willham (1963) for maternal effects, or by Willham (1972) for grand maternal effects, a fraction of the additive covariances in the MAM Eq. (11) or Eq. (16) or the GMAM Eqs. (19) or (20), should be added to the error term of the record. However, the resulting covariance matrix of error terms is still diagonal. Therefore, the mixed model equations are easily formed and solved. If the sire in the MAM (3) is inbred, then the regression coefficients in Eq. (9) should be used instead, and the fraction of the additive (co)variances will depend on the inbreeding of the sire. By a similar argument, the regression coefficients in (9) are to be used in expression (17) for the GMAM, when the maternal grandsire is inbred. Furthermore, Eqs. (21) and (22) should be used instead of Eqs. (19) and (20).

An additional issue when many dams have missing records is that estimation of σ_{Am}^2 and, especially, σ_{AoAm} becomes problematic. For example, using stochastic simulation Gerstmayr (1992) reported that estimates of σ_{Am}^2 and of σ_{AoAm} were highly inaccurate when dams were not recorded for the trait. Also, data sets having only one progeny per dam led to decreased accuracy of all estimates when compared with data sets with larger progeny groups. Similar results were obtained with a real data set by Maniatis and Pollott (2003). Notice that if phantom dams for animals with records and unknown mother identification are not included in the estimation process or the equivalent model presented here is fitted to the data, the resulting estimates are difficult to interpret as the genetic model of Willham (1963) is not correctly specified.

From a computational point of view, the gain in computing time due to avoiding the inclusion of phantom dams to predict breeding values with the mixed model equations is variable depending on the data structure and characteristics of data recording. The FORTRAN code for the direct solver used here may be simpler when fitting equivalent models with differential contributions to the coefficient matrix and residual structure, as compared to the code needed for these models when solving by iteration on data.

Conflict of interest

Appendix A. A numerical example for the MAM

In the following pedigree, letters denote unknown dams whereas numbers indicate known individuals.
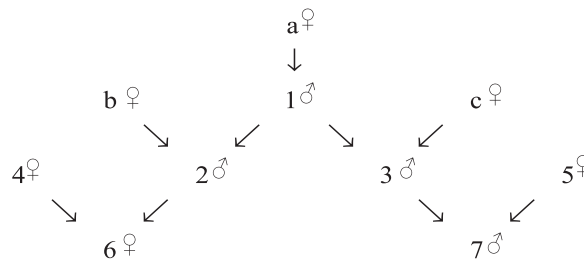
Animals 1, 2, 3, 6, and 7 have their phenotypes recorded for weaning weight in the data vector $\boldsymbol{y}' = \begin{bmatrix} 150 & 133 & 128 & 146 & 120 \end{bmatrix}$, respectively. Dams a, b, and c are "phantoms" (Westell et al., 1988; Van Vleck, 1990). A single covariate (β_1, age at weaning, measured in days) is taken as the only fixed effect, and it is related to $\boldsymbol{y}$ by the vector $\boldsymbol{x}' = \begin{bmatrix} 180 & 161 & 151 & 162 & 132 \end{bmatrix}$. Dispersion parameters are taken to betlb $\boldsymbol{G}_o = \begin{bmatrix} 100 & -25 \\ -25 & 75 \end{bmatrix}$, $\sigma_{Em}^2 = 50$ and $\sigma_{Eo}^2 = 500$.

The vector of direct breeding values is

$$\boldsymbol{a}_o' = \begin{bmatrix} a_{oa} & a_{ob} & a_{oc} & a_{o1} & a_{o2} & a_{o3} & a_{o4} & a_{o5} & a_{o6} & a_{o7} \end{bmatrix},$$

whereas the maternal breeding values are in the vector

$$\boldsymbol{a}_m' = \begin{bmatrix} a_{ma} & a_{mb} & a_{mc} & a_{m1} & a_{m2} & a_{m3} & a_{m4} & a_{m5} & a_{m6} & a_{m7} \end{bmatrix},$$

The remaining effects are those for permanent environment, and these are equal to

$$\boldsymbol{e}_m' = \begin{bmatrix} e_{ma} & e_{mb} & e_{mc} & e_{m4} & e_{m5} \end{bmatrix}$$

Respective incidence matrices for the random effects are

$$\boldsymbol{Z}_o = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}; \quad \boldsymbol{Z}_m = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}; \quad \boldsymbol{Z}_p = \boldsymbol{I}_5 \qquad (A1)$$

Let the (co)variance matrix for the breeding values, i.e. $\boldsymbol{a}' = \left[\boldsymbol{a}'_o \middle| \boldsymbol{a}'_m\right]$ be $\boldsymbol{G} = \boldsymbol{A} \otimes \boldsymbol{G}_o$, where $\boldsymbol{A}$ is the relationship matrix among the 10 animals:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & 0.50 & 0.25 & 0.25 & 0 & 0 & 0.125 & 0.125 \\ 0 & 1 & 0 & 0 & 0.50 & 0 & 0 & 0 & 0.25 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0.50 & 0 & 0 & 0 & 0.25 \\ 0.50 & 0 & 0 & 1 & 0.50 & 0.50 & 0 & 0 & 0.25 & 0.25 \\ 0.25 & 0.50 & 0 & 0.50 & 1 & 0.25 & 0 & 0 & 0.50 & 0.125 \\ 0.25 & 0 & 0.50 & 0.50 & 0.25 & 1 & 0 & 0 & 0.125 & 0.50 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0.50 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0.50 \\ 0.125 & 0.25 & 0 & 0.25 & 0.50 & 0.125 & 0.50 & 0 & 1 & 0.0625 \\ 0.125 & 0 & 0.25 & 0.25 & 0.125 & 0.50 & 0 & 0.50 & 0.0625 & 1 \end{bmatrix} \tag{A2}$$

For the vector $\boldsymbol{e}_m$ the (co)variance matrix is $\boldsymbol{I}_5 \sigma^2_{Em}$, and for the error term $\boldsymbol{R} = \boldsymbol{I}_5 \sigma^2_{Eo}$.
With all these specifications the mixed model equations are equal to:

$$\begin{bmatrix} \boldsymbol{x}'\boldsymbol{R}^{-1}\boldsymbol{x} & \boldsymbol{x}'\boldsymbol{R}^{-1}\boldsymbol{Z}_o & \boldsymbol{x}'\boldsymbol{R}^{-1}\boldsymbol{Z}_m & \boldsymbol{x}'\boldsymbol{R}^{-1}\boldsymbol{Z}_p \\ \boldsymbol{Z}'_o\boldsymbol{R}^{-1}\boldsymbol{x} & \boldsymbol{Z}'_o\boldsymbol{R}^{-1}\boldsymbol{Z}_o + \boldsymbol{A}^{-1}g^{11} & \boldsymbol{Z}'_o\boldsymbol{R}^{-1}\boldsymbol{Z}_m + \boldsymbol{A}^{-1}g^{12} & \boldsymbol{Z}'_o\boldsymbol{R}^{-1}\boldsymbol{Z}_p \\ \boldsymbol{Z}'_m\boldsymbol{R}^{-1}\boldsymbol{x} & \boldsymbol{Z}'_m\boldsymbol{R}^{-1}\boldsymbol{Z}_o + \boldsymbol{A}^{-1}g^{21} & \boldsymbol{Z}'_m\boldsymbol{R}^{-1}\boldsymbol{Z}_m + \boldsymbol{A}^{-1}g^{22} & \boldsymbol{Z}'_m\boldsymbol{R}^{-1}\boldsymbol{Z}_p \\ \boldsymbol{Z}'_p\boldsymbol{R}^{-1}\boldsymbol{x} & \boldsymbol{Z}'_p\boldsymbol{R}^{-1}\boldsymbol{Z} & \boldsymbol{Z}'_p\boldsymbol{R}^{-1}\boldsymbol{Z}_o & \boldsymbol{Z}'_p\boldsymbol{R}^{-1}\boldsymbol{Z}_p + \boldsymbol{I}_5\frac{1}{50} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\boldsymbol{a}}_o \\ \hat{\boldsymbol{a}}_m \\ \hat{\boldsymbol{e}}_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}'\boldsymbol{R}^{-1}\boldsymbol{y} \\ \boldsymbol{Z}'_o\boldsymbol{R}^{-1}\boldsymbol{y} \\ \boldsymbol{Z}'_m\boldsymbol{R}^{-1}\boldsymbol{y} \\ \boldsymbol{Z}'_p\boldsymbol{R}^{-1}\boldsymbol{y} \end{bmatrix} \tag{A3}$$

where

$$\boldsymbol{G}_0^{-1} = \begin{bmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{bmatrix}$$

By replacing with Eqs. (A1) and (A2) and the inverse of $\boldsymbol{R}$ in Eq. (A3), solutions to these equations are

$$\hat{\beta}_1 = 0.859; \quad \begin{bmatrix} \hat{a}_{oa} \\ \hat{a}_{ob} \\ \hat{a}_{oc} \\ \hat{a}_{o1} \\ \hat{a}_{o2} \\ \hat{a}_{o3} \\ \hat{a}_{o4} \\ \hat{a}_{o5} \\ \hat{a}_{o6} \\ \hat{a}_{o7} \end{bmatrix} = \begin{bmatrix} -0.191 \\ 0.053 \\ 0.174 \\ -0.637 \\ -0.440 \\ -0.125 \\ 0.255 \\ 0.242 \\ 0.419 \\ 0.543 \end{bmatrix}; \quad \begin{bmatrix} \hat{a}_{ma} \\ \hat{a}_{mb} \\ \hat{a}_{mc} \\ \hat{a}_{m1} \\ \hat{a}_{m2} \\ \hat{a}_{m3} \\ \hat{a}_{m4} \\ \hat{a}_{m5} \\ \hat{a}_{m6} \\ \hat{a}_{m7} \end{bmatrix} = \begin{bmatrix} -0.419 \\ -0.569 \\ -0.231 \\ -0.074 \\ -0.284 \\ -0.179 \\ 0.639 \\ 0.606 \\ 0.049 \\ 0.092 \end{bmatrix}; \quad \begin{bmatrix} \hat{e}_{ma} \\ \hat{e}_{mb} \\ \hat{e}_{mc} \\ \hat{e}_{m4} \\ \hat{e}_{m5} \end{bmatrix} = \begin{bmatrix} -0.339 \\ -0.404 \\ -0.136 \\ 0.511 \\ 0.485 \end{bmatrix} \tag{A4}$$

We now fit the equivalent more parsimonious MAM. The maternal breeding value of animal a is replaced by half of the maternal breeding value of 1, and those maternal breeding values of b, c, 4 and 5 are absorbed by using Eq. (12). The use of the "regressions" noticeably decreases the number of equations to be solved from 26 to 11, as the vector of breeding values for the direct effects now is $\boldsymbol{a}'_o = \begin{bmatrix} a_{o1} & a_{o2} & a_{o3} & a_{o6} & a_{o7} \end{bmatrix}$, and the vector of maternal breeding values is $\boldsymbol{a}'_m = \begin{bmatrix} a_{m1} & a_{m2} & a_{m3} & a_{m6} & a_{m7} \end{bmatrix}$. Respective incidence matrices are

$$\boldsymbol{Z}_o = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}; \quad \boldsymbol{Z}_m = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 \\ -\frac{1}{3} & \frac{2}{3} & 0 & 0 & 0 \\ -\frac{1}{3} & 0 & \frac{2}{3} & 0 & 0 \\ 0 & -\frac{1}{3} & 0 & \frac{2}{3} & 0 \\ 0 & 0 & -\frac{1}{3} & 0 & \frac{2}{3} \end{bmatrix} \tag{A5}$$

By eliminating animals a, b, c, 4 and 5, and removing the proper rows and columns in Eq. (A2) above, the resulting relationship matrix is:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0.50 & 0.50 & 0.25 & 0.25 \\ 0.50 & 1 & 0.25 & 0.50 & 0.125 \\ 0.50 & 0.25 & 1 & 0.125 & 0.50 \\ 0.25 & 0.50 & 0.125 & 1 & 0.0625 \\ 0.25 & 0.125 & 0.50 & 0.0625 & 1 \end{bmatrix} \tag{A6}$$

The covariance of the error terms is diagonal but different from an identity matrix. By using Eq. (13), we add the quantity $\frac{3}{4}(75)+50$ to the variance of the record of individual 1, whereas the use of Eq. (12) to obtain the error variance for animals 2, 3, 6 and 7 results in the quantity $\frac{2}{3}(75)+50$ added to the diagonal elements of $\boldsymbol{R}$. As a consequence, we have that

$$
\boldsymbol{R} = \begin{bmatrix} 606.25 & 0 & 0 & 0 & 0 \\ 0 & 600 & 0 & 0 & 0 \\ 0 & 0 & 600 & 0 & 0 \\ 0 & 0 & 0 & 600 & 0 \\ 0 & 0 & 0 & 0 & 600 \end{bmatrix} \tag{A7}
$$

Thus, the mixed model equations are equal to

$$
\begin{bmatrix} \boldsymbol{x}'\boldsymbol{R}^{-1}\boldsymbol{x} & \boldsymbol{x}'\boldsymbol{R}^{-1}\boldsymbol{Z}_o & \boldsymbol{x}'\boldsymbol{R}^{-1}\boldsymbol{Z}_m \\ \boldsymbol{Z}_o'\boldsymbol{R}^{-1}\boldsymbol{x} & \boldsymbol{Z}_o'\boldsymbol{R}^{-1}\boldsymbol{Z}_o+\boldsymbol{A}^{-1}g^{11} & \boldsymbol{Z}_o'\boldsymbol{R}^{-1}\boldsymbol{Z}_m+\boldsymbol{A}^{-1}g^{12} \\ \boldsymbol{Z}_m'\boldsymbol{R}^{-1}\boldsymbol{x} & \boldsymbol{Z}_m'\boldsymbol{R}^{-1}\boldsymbol{Z}_o+\boldsymbol{A}^{-1}g^{21} & \boldsymbol{Z}_m'\boldsymbol{R}^{-1}\boldsymbol{Z}_m+\boldsymbol{A}^{-1}g^{22} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\boldsymbol{a}}_o \\ \hat{\boldsymbol{a}}_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}'\boldsymbol{R}^{-1}\boldsymbol{y} \\ \boldsymbol{Z}_o'\boldsymbol{R}^{-1}\boldsymbol{y} \\ \boldsymbol{Z}_m'\boldsymbol{R}^{-1}\boldsymbol{y} \end{bmatrix} \tag{A8}
$$

The solutions of Eq. (A8) are

$$
\hat{\beta}_1 = 0.859; \quad \begin{bmatrix} \hat{a}_{o1} \\ \hat{a}_{o2} \\ \hat{a}_{o3} \\ \hat{a}_{o6} \\ \hat{a}_{o7} \end{bmatrix} = \begin{bmatrix} -0.637 \\ -0.440 \\ -0.125 \\ 0.419 \\ 0.543 \end{bmatrix}; \quad \begin{bmatrix} \hat{a}_{m1} \\ \hat{a}_{m2} \\ \hat{a}_{m3} \\ \hat{a}_{m6} \\ \hat{a}_{m7} \end{bmatrix} = \begin{bmatrix} -0.074 \\ -0.284 \\ -0.179 \\ 0.049 \\ 0.092 \end{bmatrix}; \tag{A9}
$$

Clearly, solutions of the fixed effect and the breeding values that are alike in Eq. (A9) and Eq. (A4) are equal. The equivalence between models is further observed while comparing the estimated expectation and the covariance matrix of $\boldsymbol{y}$ under both models, which are respectively equal to

$$
E(\boldsymbol{y}) = \boldsymbol{x}\beta_1 = \begin{bmatrix} 180 \\ 161 \\ 151 \\ 162 \\ 132 \end{bmatrix} 0.859 = \begin{bmatrix} 154.62 \\ 138.30 \\ 129.71 \\ 139.16 \\ 113.39 \end{bmatrix} \tag{A10}
$$

and

$$
\boldsymbol{V} = \begin{bmatrix} 700 & 43.75 & 43.75 & 21.875 & 21.875 \\ 43.75 & 700 & 25 & 43.75 & 12.50 \\ 43.75 & 25 & 700 & 12.50 & 43.75 \\ 21.875 & 43.75 & 12.50 & 700 & 6.25 \\ 21.875 & 12.50 & 43.75 & 6.25 & 700 \end{bmatrix} \tag{A11}
$$

Expression (A11) was calculated by using

$$
V = Z_o A Z_o' 100 - 25(Z_o A Z_m' + Z_m A Z_o') + Z_m A Z_m' 75 + Z_p Z_p' 50 + I_5\,500
$$

for the conventional MAM; and by taking

$$
V = Z_o A Z_o' 100 - 25(Z_o A Z_m' + Z_m A Z_o') + Z_m A Z_m' 75 + R,
$$

for the equivalent model.

Appendix B. Derivation of the variance of $\varepsilon_{o\,i}$ for the model with grand maternal effects and missing grand dams

Case 1. *Maternal grand dam unknown and maternal grandsire known.*

In the following pedigree, l represents the maternal grandsire of the animal with record i, g is the maternal grand dam of i, and j is the dam of i.

$$
\begin{array}{ccc}
\male\, l & & g\,\female \\
& \searrow \quad \swarrow & \\
& j \;\; \female & \\
& \downarrow & \\
& i &
\end{array}
$$
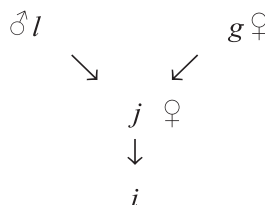
Starting with the following equation for the GMAM

$$y_{ijg} = x_i'\beta + a_{o\,i} + a_{m\,j} + a_{n\,g} + e_{m\,j} + e_{o\,i} \tag{B1}$$

The variance of an observation is equal to

$$
\begin{aligned}
Var\left(y_{ijg}\right) &= Var(a_{oi} + a_{mj} + a_{ng} + e_{mj} + e_{oi}) \\
&= Var(a_{oi}) + Var(a_{mj}) + Var(a_{ng}) + 2\text{cov}(a_{oi}, a_{mj}) \\
&\quad + 2\text{cov}(a_{oi}, a_{ng}) + 2\text{cov}(a_{mj}, a_{ng}) + Var(e_{mj}) + Var(e_{oi}) \\
&= (1+F_i)\sigma_{Ao}^2 + (1+F_j)\sigma_{Am}^2 + (1+F_g)\sigma_{An}^2 + 2\left[F_i + \tfrac{1}{2}(1+F_j)\right]\sigma_{AoAm} \\
&\quad + 2\left[F_i + \tfrac{1}{4}(1+F_g)\right]\sigma_{AoAn} + 2\left[F_j + \tfrac{1}{2}(1+F_g)\right]\sigma_{AmAn} + \sigma_{Em}^2 + \sigma_{Eo}^2
\end{aligned} \tag{B2}
$$

When the maternal grand dam g is unknown, it may be safely assume in Eq. (B2) that $F_g = F_j = 0$. Moreover, the grand maternal breeding value of the grand dam can be replaced by the linear function of the grand maternal breeding values of its daughter (j) and the sire of its daughter (l) displayed in Eq. (17) so that:

$$y_{ijl} = x_i'\beta + a_{oi} + a_{mj} + \tfrac{2}{3}a_{nj} - \tfrac{1}{3}a_{nl} + e_{mj} + e_{oi} \tag{B3}$$

On using the variance operator in (B3) results in

$$
\begin{aligned}
Var\left(y_{ijl}\right) &= Var\left(a_{oi} + a_{mj} + \tfrac{2}{3}a_{nj} - \tfrac{1}{3}a_{nl} + e_{mj} + \varepsilon_{oi}\right) = Var(a_{oi}) + Var(a_{mj}) + \tfrac{4}{9}Var(a_{nj}) + \tfrac{1}{9}Var(a_{nl}) \\
&\quad + Var(e_{mj}) + Var(\varepsilon_{oi}) + 2\text{cov}(a_{oi}, a_{mj}) + 2\text{cov}(a_{oi}, \tfrac{2}{3}a_{nj}) + 2\text{cov}(a_{oi}, -\tfrac{1}{3}a_{nl}) + 2\text{cov}(a_{mj}, \tfrac{2}{3}a_{nj}) \\
&\quad + 2\text{cov}(a_{mj}, -\tfrac{1}{3}a_{nl}) + 2\text{cov}(\tfrac{2}{3}a_{nj}, -\tfrac{1}{3}a_{nl}) = (1+F_i)\sigma_{Ao}^2 + (1+F_j)\sigma_{Am}^2 + \tfrac{4}{9}(1+F_j)\sigma_{An}^2 \\
&\quad + \tfrac{1}{9}(1+F_l)\sigma_{An}^2 + \sigma_{Em}^2 + Var(\varepsilon_{oi}) + 2(\tfrac{1}{2})\sigma_{AoAm} + 2(\tfrac{2}{3})\tfrac{1}{2}\sigma_{AoAn} + 2(-\tfrac{1}{3})\tfrac{1}{4}\sigma_{AoAn} + 2(\tfrac{2}{3})(1+F_j)\sigma_{AmAn} \\
&\quad + 2(-\tfrac{1}{3})\tfrac{1}{2}\sigma_{AmAn} + 2\tfrac{2}{3}(-\tfrac{1}{3})\tfrac{1}{2}\sigma_{An}^2 = (1+F_i)\sigma_{Ao}^2 + (1+F_j)\sigma_{Am}^2 + \tfrac{4}{9}(1+F_j)\sigma_{An}^2 \\
&\quad + \tfrac{1}{9}(1+F_l)\sigma_{An}^2 + \sigma_{Em}^2 + Var(\varepsilon_{oi}) + \sigma_{AoAm} + \tfrac{2}{3}\sigma_{AoAn} - \tfrac{1}{6}\sigma_{AoAn} + \tfrac{4}{3}(1+F_j)\sigma_{AmAn} - \tfrac{1}{3}\sigma_{AmAn} - \tfrac{2}{9}\sigma_{An}^2 = (1+F_i)\sigma_{Ao}^2 \\
&\quad + (1+F_j)\sigma_{Am}^2 + \left[\tfrac{4}{9}(1+F_j) + \tfrac{1}{9}(1+F_l) - \tfrac{2}{9}\right]\sigma_{An}^2 + \sigma_{Em}^2 + Var(\varepsilon_{oi}) + \sigma_{AoAm} + \left[\tfrac{2}{3} - \tfrac{1}{6}\right]\sigma_{AoAn} + \left[\tfrac{4}{3}(1+F_j) - \tfrac{1}{3}\right]\sigma_{AmAn} \\
&= (1+F_i)\sigma_{Ao}^2 + (1+F_j)\sigma_{Am}^2 + \left[\tfrac{1}{3} + \tfrac{(4F_j + F_l)}{9}\right]\sigma_{An}^2 + \sigma_{Em}^2 + Var(\varepsilon_{oi}) + \sigma_{AoAm} + \tfrac{1}{2}\sigma_{AoAn} + (1+\tfrac{4}{3}F_j)\sigma_{AmAn}
\end{aligned}
$$

As g is unknown, it is reasonable to assume that the dam of i is not inbred, i.e. $F_j = 0$. Consequently, on using this result in the previous expression produces the variance of a record under the GMAM when the grand dam is missing as follows

$$Var\left(y_{ijl}\right) = (1+F_i)\sigma_{Ao}^2 + \sigma_{Am}^2 + \left[\tfrac{1}{3} + \tfrac{F_l}{9}\right]\sigma_{An}^2 + \sigma_{Em}^2 + Var(\varepsilon_{oi}) + \sigma_{AoAm} + \tfrac{1}{2}\sigma_{AoAn} + \sigma_{AmAn} \tag{B4}$$

The fraction of the additive variance that is not accounted by the model (B3) is thus obtained by subtracting Eq. (B4) from (B2), and then solving for Var(ε_{oi}). In case the maternal grandsire is not inbred (i.e., $F_l = 0$), the residual variance is equal to

$$Var(\varepsilon_{oi}) = \sigma_{Eo}^2 + \tfrac{2}{3}\sigma_{An}^2 \tag{B5}$$

Case 2. *Maternal grand dam and maternal grandsire are both unknown.*

When the maternal grandsire of the animal with record is also unknown, an equivalent GMAM consistent with the specification of the covariance structure given by Willham (1972) is equal to

$$y_{ij} = \mathbf{x}_i'\boldsymbol{\beta} + a_{oi} + a_{mj} + \tfrac{1}{2}a_{nj} + e_{mj} + \varepsilon_{oi} \tag{B6}$$

Dam j has both parents unknown, so that $F_k = F_l = F_j = 0$. Now, taking the variance operator on (B6) produces

$$
\begin{aligned}
Var\left(y_{ij}\right) &= Var\left(a_{oi} + a_{mj} + \tfrac{1}{2}a_{nj} + e_{mj} + \varepsilon_{oi}\right) = Var(a_{oi}) + Var(a_{mj}) + \tfrac{1}{4}Var(a_{nj}) + Var(e_{mj}) + Var(\varepsilon_{oi}) \\
&\quad + 2\text{cov}(a_{oi}, a_{mj}) + 2\text{cov}(a_{oi}, \tfrac{1}{2}a_{nj}) + 2\text{cov}(a_{mj}, \tfrac{1}{2}a_{nj}) = (1+F_i)\sigma_{Ao}^2 + \sigma_{Am}^2 + \tfrac{1}{4}\sigma_{An}^2 + \sigma_{Em}^2 \\
&\quad + Var(\varepsilon_{oi}) + 2(\tfrac{1}{2})\sigma_{AoAm} + 2(\tfrac{1}{4})\sigma_{AoAn} + 2(\tfrac{1}{2})\sigma_{AmAn}
\end{aligned}
$$

Therefore

$$Var\left(y_{ij}\right) = (1+F_i)\sigma_{Ao}^2 + \sigma_{Am}^2 + \tfrac{1}{4}\sigma_{An}^2 + \sigma_{Em}^2 + Var(\varepsilon_{oi}) + \sigma_{AoAm} + \tfrac{1}{2}\sigma_{AoAm} + \sigma_{AmAn} \tag{B7}$$
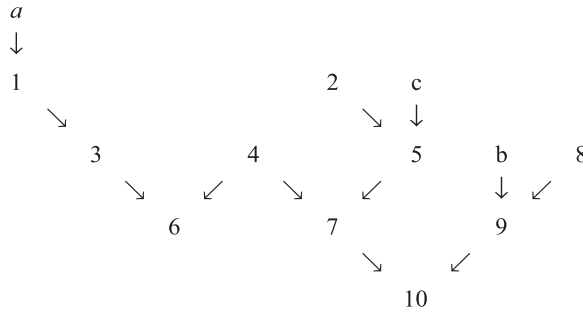
And, on subtracting (B7) to (B2) we end up with the following variance of the error term

$$Var(\varepsilon_{oi}) = \sigma_{Eo}^2 + \tfrac{3}{4}\sigma_{An}^2 \tag{B8}$$

Expressions (B5) and (B8) display the heterogeneity of residual variance due to missing pedigree information in the GMAM.

Appendix C. A numerical example of the grand maternal animal model with missing maternal grand dams

The pedigree for this example is seen below.

Dams a, b and c are unknown maternal grand dams ("phantom"). Individuals 3, 6, 7 and 10 have weaning weight records in $\mathbf{y}$ such that $\mathbf{y}' = \begin{bmatrix} 150 & 133 & 128 & 146 \end{bmatrix}$. Again, age at weaning in days is the only fixed effect (β_1), and is related to $\mathbf{y}$ by

the incidence vector $\mathbf{x}' = \begin{bmatrix} 180 & 161 & 151 & 162 \end{bmatrix}$. Covariance components are $\mathbf{G}_o = \begin{bmatrix} 100 & -25 & 15 \\ -25 & 75 & 18 \\ 15 & 18 & 60 \end{bmatrix}$, $\sigma^2_{Em} = 50$, and

$\sigma^2_{Eo} = 500$. The fitting of the GMAM of Willham (1972) necessitates the following vectors of direct, maternal and grand maternal breeding values, and permanent maternal environmental effects:

$$\mathbf{a}'_o = \begin{bmatrix} a_{oa} & a_{ob} & a_{oc} & a_{o1} & a_{o2} & a_{o3} & a_{o4} & a_{o5} & a_{o6} & a_{o7} & a_{o8} & a_{o9} & a_{o10} \end{bmatrix};$$

$$\mathbf{a}'_m = \begin{bmatrix} a_{ma} & a_{mb} & a_{mc} & a_{m1} & a_{m2} & a_{m3} & a_{m4} & a_{m5} & a_{m6} & a_{m7} & a_{m8} & a_{m9} & a_{m10} \end{bmatrix};$$

$$\mathbf{a}'_n = \begin{bmatrix} a_{na} & a_{nb} & a_{nc} & a_{n1} & a_{n2} & a_{n3} & a_{n4} & a_{n5} & a_{n6} & a_{n7} & a_{n8} & a_{n9} & a_{n10} \end{bmatrix};$$

$$\mathbf{e}'_m = \begin{bmatrix} e_{m1} & e_{m3} & e_{m5} & e_{m9} \end{bmatrix};$$

The respective incidence matrices of those effects are

$$\mathbf{Z}_o = \begin{bmatrix} 0&0&0&0&0&1&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&1 \end{bmatrix}; \quad \mathbf{Z}_m = \begin{bmatrix} 0&0&0&1&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&1&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&1&0 \end{bmatrix};$$

$$\mathbf{Z}_n = \begin{bmatrix} 1&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0&0&0&0 \end{bmatrix}; \quad \mathbf{Z}_p = \mathbf{I}_4; \tag{C1}$$

The relationship matrix $\mathbf{A}$ is

$$\mathbf{A} = \begin{bmatrix}
1 & 0 & 0 & 0.50 & 0 & 0.25 & 0 & 0 & 0.125 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.50 & 0.25 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0.50 & 0 & 0.25 & 0 & 0 & 0.125 \\
0.50 & 0 & 0 & 1 & 0 & 0.50 & 0 & 0 & 0.25 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0.50 & 0 & 0.25 & 0 & 0 & 0.125 \\
0.25 & 0 & 0 & 0.50 & 0 & 1 & 0 & 0 & 0.50 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0.50 & 0.50 & 0 & 0 & 0.25 \\
0 & 0 & 0.50 & 0 & 0.50 & 0 & 0 & 1 & 0 & 0.50 & 0 & 0 & 0.25 \\
0.125 & 0 & 0 & 0.25 & 0 & 0.50 & 0.50 & 0 & 1 & 0.25 & 0 & 0 & 0.125 \\
0 & 0 & 0.25 & 0 & 0.25 & 0 & 0.50 & 0.50 & 0.25 & 1 & 0 & 0 & 0.50 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0.50 & 0.25 \\
0 & 0.50 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.50 & 1 & 0.50 \\
0 & 0.25 & 0.125 & 0 & 0.125 & 0 & 0.25 & 0.25 & 0.125 & 0.50 & 0.25 & 0.50 & 1
\end{bmatrix} \tag{C2}$$

The covariance matrix for the environmental maternal effects is $Var(\mathbf{e}_m) = \mathbf{I}_4 \sigma^2_{Em}$, and for the error vector is $\mathbf{R} = \mathbf{I}_4 \sigma^2_{Eo}$.

With all these specifications, MME are equal to

$$
\begin{bmatrix}
\mathbf{x'R^{-1}x} & \mathbf{x'R^{-1}Z_o} & \mathbf{x'R^{-1}Z_m} & \mathbf{x'R^{-1}Z_n} & \mathbf{x'R^{-1}Z_p} \\
\mathbf{Z'_oR^{-1}x} & \mathbf{Z'_oR^{-1}Z_o}+\mathbf{A^{-1}}g^{11} & \mathbf{Z'_oR^{-1}Z_m}+\mathbf{A^{-1}}g^{12} & \mathbf{Z'_oR^{-1}Z_n}+\mathbf{A^{-1}}g^{13} & \mathbf{Z'_oR^{-1}Z_p} \\
\mathbf{Z'_mR^{-1}x} & \mathbf{Z'_mR^{-1}Z_o}+\mathbf{A^{-1}}g^{21} & \mathbf{Z'_mR^{-1}Z_m}+\mathbf{A^{-1}}g^{22} & \mathbf{Z'_mR^{-1}Z_n}+\mathbf{A^{-1}}g^{23} & \mathbf{Z'_mR^{-1}Z_p} \\
\mathbf{Z'_nR^{-1}x} & \mathbf{Z'_nR^{-1}Z_o}+\mathbf{A^{-1}}g^{31} & \mathbf{Z'_nR^{-1}Z_m}+\mathbf{A^{-1}}g^{32} & \mathbf{Z'_nR^{-1}Z_n}+\mathbf{A^{-1}}g^{33} & \mathbf{Z'_nR^{-1}Z_p} \\
\mathbf{Z'_pR^{-1}x} & \mathbf{Z'_pR^{-1}Z_o} & \mathbf{Z'_pR^{-1}Z_m} & \mathbf{Z'_pR^{-1}Z_n} & \mathbf{Z'_p{}'R^{-1}Z_p}+\mathbf{I_4}\tfrac{1}{50}
\end{bmatrix}
\begin{bmatrix}
\hat{\beta}_1 \\ \hat{\mathbf{a}}_o \\ \hat{\mathbf{a}}_m \\ \hat{\mathbf{a}}_n \\ \hat{\mathbf{e}}_m
\end{bmatrix}
=
\begin{bmatrix}
\mathbf{x'R^{-1}y} \\ \mathbf{Z'_oR^{-1}y} \\ \mathbf{Z'_mR^{-1}y} \\ \mathbf{Z'_oR^{-1}y} \\ \mathbf{Z'_pR^{-1}y}
\end{bmatrix}
\quad (C3)
$$

where

$$
\mathbf{G_0^{-1}} =
\begin{bmatrix}
g^{11} & g^{12} & g^{13} \\
g^{21} & g^{22} & g^{23} \\
g^{31} & g^{32} & g^{33}
\end{bmatrix}
$$

After replacing with Eqs. (C1) and (C2) and the inverse of $\mathbf{R}$, the solutions of the system are

$$
\hat{\beta}_1 = 0.852; \quad
\begin{bmatrix}
\hat{a}_{oa} \\ \hat{a}_{ob} \\ \hat{a}_{oc} \\ \hat{a}_{o1} \\ \hat{a}_{o2} \\ \hat{a}_{o3} \\ \hat{a}_{o4} \\ \hat{a}_{o5} \\ \hat{a}_{o6} \\ \hat{a}_{o7} \\ \hat{a}_{o8} \\ \hat{a}_{o9} \\ \hat{a}_{o10}
\end{bmatrix}
=
\begin{bmatrix}
-0.168 \\ 0.281 \\ 0.908 \\ -0.254 \\ 0.110 \\ -0.492 \\ -0.060 \\ 0.212 \\ -0.525 \\ 0.264 \\ 0.128 \\ 0.332 \\ 0.809
\end{bmatrix}; \quad
\begin{bmatrix}
\hat{a}_{ma} \\ \hat{a}_{mb} \\ \hat{a}_{mc} \\ \hat{a}_{m1} \\ \hat{a}_{m2} \\ \hat{a}_{m3} \\ \hat{a}_{m4} \\ \hat{a}_{m5} \\ \hat{a}_{m6} \\ \hat{a}_{m7} \\ \hat{a}_{m8} \\ \hat{a}_{m9} \\ \hat{a}_{m10}
\end{bmatrix}
=
\begin{bmatrix}
-0.301 \\ -0.504 \\ -0.098 \\ -0.504 \\ -0.074 \\ -0.417 \\ 0.015 \\ -0.160 \\ -0.139 \\ -0.119 \\ 0.319 \\ 0.731 \\ 0.178
\end{bmatrix}; \quad
\begin{bmatrix}
\hat{a}_{na} \\ \hat{a}_{nb} \\ \hat{a}_{nc} \\ \hat{a}_{n1} \\ \hat{a}_{n2} \\ \hat{a}_{n3} \\ \hat{a}_{n4} \\ \hat{a}_{n5} \\ \hat{a}_{n6} \\ \hat{a}_{n7} \\ \hat{a}_{n8} \\ \hat{a}_{n9} \\ \hat{a}_{n10}
\end{bmatrix}
=
\begin{bmatrix}
-0.444 \\ 0.744 \\ -0.079 \\ -0.563 \\ 0.002 \\ -0.417 \\ 0.009 \\ -0.036 \\ -0.250 \\ 0.005 \\ 0.130 \\ 0.567 \\ 0.363
\end{bmatrix}; \quad
\begin{bmatrix}
\hat{e}_{m1} \\ \hat{e}_{m3} \\ \hat{e}_{m5} \\ \hat{e}_{m9}
\end{bmatrix}
=
\begin{bmatrix}
-0.181 \\ -0.248 \\ -0.067 \\ 0.511
\end{bmatrix}
\quad (C4)
$$

As in MAM (3), fitting of models (23) and (24) allows a strong reduction in the number of equations. In this case, the vectors $\mathbf{a}_o$, $\mathbf{a}_m$ and $\mathbf{e}_m$ are equal to

$$
\mathbf{a'_o} = \begin{bmatrix} a_{o1} & a_{o2} & a_{o3} & a_{o4} & a_{o5} & a_{o6} & a_{o7} & a_{o8} & a_{o9} & a_{o10} \end{bmatrix};
$$

$$
\mathbf{a'_m} = \begin{bmatrix} a_{m1} & a_{m2} & a_{m3} & a_{m4} & a_{m5} & a_{m6} & a_{m7} & a_{m8} & a_{m9} & a_{m10} \end{bmatrix};
$$

$$
\mathbf{a'_n} = \begin{bmatrix} a_{n1} & a_{n2} & a_{n3} & a_{n4} & a_{n5} & a_{n6} & a_{n7} & a_{n8} & a_{n9} & a_{n10} \end{bmatrix};
$$

$$
\mathbf{e'_m} = \begin{bmatrix} e_{m1} & e_{m3} & e_{m5} & e_{m9} \end{bmatrix};
$$

No breeding value for any type of effects was included for missing maternal. In matrix $\mathbf{Z}_n$, the grand maternal breeding value of animal 3 was replaced by $1/2 a_{n1}$, the one for 7 by $-1/3 a_{n2}+2/3 a_{n5}$, and for animal 10 the quantity added was $-1/3 a_{n8}+2/3 a_{n9}$. All in all, incidence matrices are equal to

$$
\mathbf{Z}_o =
\begin{bmatrix}
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}; \quad
\mathbf{Z}_m =
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix};
$$

$$
\mathbf{Z}_n =
\begin{bmatrix}
\tfrac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -\tfrac{1}{3} & 0 & 0 & \tfrac{2}{3} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -\tfrac{1}{3} & \tfrac{2}{3} & 0
\end{bmatrix}; \quad
\mathbf{Z}_p = \mathbf{I}_4
\quad (C5)
$$

Now matrix $\boldsymbol{A}$ is

$$
\boldsymbol{A} = \begin{bmatrix}
1 & 0 & 0.50 & 0 & 0 & 0.25 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0.50 & 0 & 0.25 & 0 & 0 & 0.125 \\
0.50 & 0 & 1 & 0 & 0 & 0.50 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0.50 & 0.50 & 0 & 0 & 0.25 \\
0 & 0.50 & 0 & 0 & 1 & 0 & 0.50 & 0 & 0 & 0.25 \\
0.25 & 0 & 0.50 & 0.50 & 0 & 1 & 0.25 & 0 & 0 & 0.125 \\
0 & 0.25 & 0 & 0.50 & 0.50 & 0.25 & 1 & 0 & 0 & 0.50 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0.50 & 0.25 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.50 & 1 & 0.50 \\
0 & 0.125 & 0 & 0.25 & 0.25 & 0.125 & 0.50 & 0.25 & 0.50 & 1
\end{bmatrix} \tag{C6}
$$

Finally, $Var(\boldsymbol{e}_m) = \boldsymbol{I}_4 \sigma_{Em}^2$ and

$$
\boldsymbol{R} = \begin{bmatrix}
545 & 0 & 0 & 0 \\
0 & 500 & 0 & 0 \\
0 & 0 & 540 & 0 \\
0 & 0 & 0 & 540
\end{bmatrix} \tag{C7}
$$

In Eq. (C7), 3/4(60) was added to the error variance (500) for animal *1*, whereas 2/3(60) was added to those diagonal elements of individuals 7 and 10. On using Eqs. (C5) to (C7) in Eq. (C3), we obtain the following solutions

$$
\hat{\beta}_1 = 0.852; \quad
\begin{bmatrix}
\hat{a}_{o1} \\ \hat{a}_{o2} \\ \hat{a}_{o3} \\ \hat{a}_{o4} \\ \hat{a}_{o5} \\ \hat{a}_{o6} \\ \hat{a}_{o7} \\ \hat{a}_{o8} \\ \hat{a}_{o9} \\ \hat{a}_{o10}
\end{bmatrix} =
\begin{bmatrix}
-0.254 \\ 0.110 \\ -0.492 \\ -0.060 \\ 0.212 \\ -0.525 \\ 0.264 \\ 0.128 \\ 0.332 \\ 0.809
\end{bmatrix}; \quad
\begin{bmatrix}
\hat{a}_{m1} \\ \hat{a}_{m2} \\ \hat{a}_{m3} \\ \hat{a}_{m4} \\ \hat{a}_{m5} \\ \hat{a}_{m6} \\ \hat{a}_{m7} \\ \hat{a}_{m8} \\ \hat{a}_{m9} \\ \hat{a}_{m10}
\end{bmatrix} =
\begin{bmatrix}
-0.504 \\ -0.074 \\ -0.417 \\ 0.015 \\ -0.160 \\ -0.139 \\ -0.119 \\ 0.319 \\ 0.731 \\ 0.178
\end{bmatrix}; \quad
\begin{bmatrix}
\hat{a}_{n1} \\ \hat{a}_{n2} \\ \hat{a}_{n3} \\ \hat{a}_{n4} \\ \hat{a}_{n5} \\ \hat{a}_{n6} \\ \hat{a}_{n7} \\ \hat{a}_{n8} \\ \hat{a}_{n9} \\ \hat{a}_{n10}
\end{bmatrix} =
\begin{bmatrix}
-0.563 \\ 0.002 \\ -0.417 \\ 0.009 \\ -0.036 \\ -0.250 \\ 0.005 \\ 0.130 \\ 0.567 \\ 0.363
\end{bmatrix}; \quad
\begin{bmatrix}
\hat{e}_{m1} \\ \hat{e}_{m3} \\ \hat{e}_{m5} \\ \hat{e}_{m9}
\end{bmatrix} =
\begin{bmatrix}
-0.181 \\ -0.248 \\ -0.067 \\ 0.511
\end{bmatrix} \tag{C8}
$$

Solutions for all animals included in both systems of MME and displayed in Eq. (C4) and Eq. (C8), are equal. Model equivalence is further observed when comparing the estimated expectation of $\boldsymbol{y}$ under both models, which is equal to

$$
E(\boldsymbol{y}) = \boldsymbol{x}\beta_1 = \begin{bmatrix} 180 \\ 161 \\ 151 \\ 162 \end{bmatrix} 0.852 = \begin{bmatrix} 153.36 \\ 137.17 \\ 128.65 \\ 138.02 \end{bmatrix} \tag{C9}
$$

Calculating the covariance matrix under both models as

$$
\boldsymbol{V} = \boldsymbol{Z}_o\boldsymbol{A}\boldsymbol{Z}_o'100 + \boldsymbol{Z}_m\boldsymbol{A}\boldsymbol{Z}_m'75 + \boldsymbol{Z}_n\boldsymbol{A}\boldsymbol{Z}_n'60 - 25\left(\boldsymbol{Z}_o\boldsymbol{A}\boldsymbol{Z}_m' + \boldsymbol{Z}_m\boldsymbol{A}\boldsymbol{Z}_o'\right) + 15\left(\boldsymbol{Z}_o\boldsymbol{A}\boldsymbol{Z}_n' + \boldsymbol{Z}_n\boldsymbol{A}\boldsymbol{Z}_o'\right) + 18\left(\boldsymbol{Z}_m\boldsymbol{A}\boldsymbol{Z}_n' + \boldsymbol{Z}_n\boldsymbol{A}\boldsymbol{Z}_m'\right) + \boldsymbol{Z}_p\boldsymbol{Z}_p'50 + \boldsymbol{R} \tag{C10}
$$

results in

$$
\boldsymbol{V} = \begin{bmatrix}
785.50 & 118.125 & 0 & 0 \\
118.125 & 785.50 & 25 & 12.50 \\
0 & 25 & 785.50 & 45.625 \\
0 & 12.50 & 45.625 & 785.50
\end{bmatrix} \tag{C11}
$$

References

Bulmer, M.G., 1985. The Mathematical Theory of Quantitative Genetics, 2nd ed. Clarendon Press, Oxford, United Kingdom.

Cantet, R.J.C., Schaeffer, L.R., Smith, C., 1992. Reduced animal model with differential genetic grouping for direct and maternal effects. J. Anim. Sci. 70, 1730–1741.

Dodenhoff, J., Van Vleck, L.D., Kachman, S.D., Koch, R.M., 1998. Parameter estimates for direct, maternal and grand maternal genetic effects for birth weight and weaning weight in Hereford cattle. J. Anim. Sci. 76, 2521–2527.

Dodenhoff, J., Van Vleck, L.D., Gregory, K.E., 1999a. Estimation of direct, maternal, and grandmaternal genetic effects for weaning weight in several breeds of beef cattle. J. Anim. Sci. 77, 840–845.

Dodenhoff, J., Van Vleck, L.D., Wilson, D.E., 1999b. Comparison of models to estimate genetic effects for weaning weight of Angus cattle. J. Anim. Sci. 77, 3176–3184.

Gerstmayr, S., 1992. Impact of the data structure on the reliability of the estimated genetic parameters in an animal model with maternal effects. J. Anim. Breed. Genet. 109, 321–336.

Henderson, C.R., 1984. Applications of Linear Models in Animal Breeding. University of Guelph, Guelph.

Henderson, C.R., 1985. Equivalent linear models to reduce computations. J. Dairy Sci. 68, 2267–2277.

Maniatis, N., Pollott, G.E., 2003. The impact of data structure on genetic (co)variance components of early growth in sheep, estimated using an animal model with maternal effects. J. Anim. Sci. 81, 101–108.

Schaeffer, L.R., Henderson, C.R., 1983. Best linear unbiased prediction when error vector is correlated with other random vectors in the model. Genet. Sel. Evol. 15, 395–400.

Van Vleck, L.D., 1990. Breeding value prediction with maternal genetic groups. J. Anim. Sci. 58, 3998–4013.

Westell, R.A., Quaas, R.L., Van Vleck, L.D., 1988. Genetic groups in animal model. J. Dairy Sci. 69, 1310–1318.

Willham, R.L., 1963. The covariance between relatives for characters composed of components contributed by related individuals. Biometrics 19, 18–27.

Willham, R.L., 1972. The role of maternal effects in animal breeding: III. Biometrical aspects of maternal effects in animals. J. Anim. Sci. 35, 1288–1293.