Reduced Animal Model with Differential Genetic Grouping for Direct and Maternal Effects[1]

R.J.C. Cantet*,†,[2], L. R. Schaeffer*, and C. Smith*

*Centre for Genetic Improvement of Livestock, Department of Animal and Poultry Science, University of Guelph, Guelph, Ontario N1G 2W1, Canada and †Department of Animal Science, University of Illinois, Urbana 61801

ABSTRACT: Mixed-model equations for the reduced animal model with maternal effects and different genetic grouping of unknown parents for additive direct and maternal effects are derived. The matrices that relate the expected value and the variance of the breeding values of non-parents to the parents, as well as the different contributions of parental and non-parental breeding values, to the resulting mixed-model equations are presented. Mis-specification of additive maternal variance and the additive covariance between direct and maternal effects, arising from missing information on the dams of known individuals with records, is discussed. To avoid an incorrect specification of the variance-covariance matrix of the records without having to invert a nondiagonal variance of the residual terms, the breeding values of the unknown dams of individuals with records are included in the equations. Breeding values of non-parents are back-solved after the solutions for genetic groups and breeding values of parents are computed as simply as in cases in which maternal effects are absent. A numerical example is included to illustrate the derivations.

Key Words: Animal Models, Genetic Analysis, Maternal Effects

J. Anim. Sci. 1992. 70:1730–1741

Introduction

Maternal effects are currently being included in the genetic evaluation of animals for breeding in beef cattle populations (Benyshek et al., 1988). Best linear unbiased prediction (Henderson, 1973) under the animal model (**AM**; Henderson and Quaas, 1976) and reduced animal model (**RAM**; Quaas and Pollak, 1980) are used for this purpose. A large system of equations must be solved to use AM, but often not all solutions are needed for selection and mating decisions. This is especially true for animals that are not parents, and these usually outnumber parent animals.

Modifications of the Animal Model

The RAM is an equivalent model, in the sense described by Henderson (1985), to the AM, but fewer equations must be solved. Solutions for RAM are obtained by absorbing non-parents, or by modifying the procedure for constructing Henderson's mixed-model equations (**MME**) by expressing non-parental breeding values (**BV**) in terms of parental BV.

Genetic grouping (Robinson, 1986; Quaas, 1988; Westell et al., 1988) is a means of treating incomplete pedigree information to account for the effects of past selection on estimators and predictors. Westell et al. (1988) gave a set of rules for incorporating genetic grouping in the AM. These are analogous to the simple rules needed for calculating the inverse of an additive relationship

[1]This work was supported by the Natural Sciences and Engineering Research Council of Canada and the University of Illinois. R. L. Quaas critically reviewed the original manuscript, resulting in a substantially improved version. Any mistake is the responsibility of the first author.

[2]On leave from Departamento de Zootecnia, Facultad de Agronomía, Universidad de Buenos Aires, Av. San Martin 4453, (1417) Capital Federal, Argentina. To whom correspondence should be addressed.

Received March 8, 1991.
Accepted December 21, 1991.

matrix (Henderson, 1976). However, for large data sets, the AM may be computationally unfeasible, especially if maternal effects are included in the model.

Quaas (1988) found an explicit formula for the matrix ($\boldsymbol{Q}$) that relates individuals to their populations means: the genetic groups. This formula involves a function (more precisely, $[\boldsymbol{I} - \boldsymbol{P}]^{-1}$) of the matrix $\boldsymbol{P}$ that describes the passage of genes from parents to offspring. Further, Quaas (1988) showed that the vector of BV ($\boldsymbol{a}$) and the matrix of additive relationships among individuals ($\boldsymbol{A}$) are also functions of $(\boldsymbol{I} - \boldsymbol{P})^{-1}$. The fact the $(\boldsymbol{I} - \boldsymbol{P})^{-1}$ also describes the flow of genes from ancestor to offspring (Henderson, 1976; Quaas, 1988) and that this matrix appears in $\mathbf{a}$, $\boldsymbol{A}$, and $\boldsymbol{Q}$, highlights that genetic grouping via the matrix $\boldsymbol{Q}$ is consistent with the rules of additive inheritance.

Van Vleck (1990a) extended genetic grouping theory to accommodate maternal effects. His formulation requires that the same criterion be used to assign groups for both direct and maternal effects. Further, Van Vleck (1990b) obtained RAM equations for this model. However, different genetic trends for direct and maternal effects on weaning weight have been reported by Benyshek et al. (1988) and Cantet (1990). In this situation, assigning different groups may avoid confounding between direct and maternal groups and with other effects in the model. Cantet (1990) extended genetic grouping theory to accommodate differential criteria when assigning groups for direct and maternal effects.

An important issue when working with AM or RAM for a maternally influenced trait is that misspecification of genetic covariance arises when identification on dams of individuals in the pedigree is missing. Calculating additive genetic covariances between any two individuals when maternal effects are present requires knowledge of the two individuals and their dams (Willham, 1963). The objective of this study is to provide a RAM with genetic grouping and with maternal effects that is an alternative to the one presented by Van Vleck (1990b). The equations obtained include missing dams of individuals with records and allow the use of different criteria to assign groups for direct and for maternal effects. The presentation is based on the concepts described by Quaas (1988) and complements the research described by Cantet (1990). Although several matrices are employed to derive the MME for RAM with genetic grouping, most of them are not explicitly needed to perform the actual computations. For the purpose of notation, a letter "o" used as a subscript will refer to direct effects, where "m" as a subscript will refer to maternal effects. The capital letters "P" and "N", used either as superscripts or subscripts, will refer to parent and nonparent, respectively. The subscript "b" will be used for the unknown, missing, or "phantom" (Westell et al., 1988) parents of "known" individuals with or without records. The use of b (for "base") is to keep consistency with the notation of Quaas (1988) and does not imply that all missing parents are treated as belonging to the same "base" population (Taylor and Tomaszewski, 1989).

Animal Model with Genetic Groups

An AM with genetic grouping that includes maternal effects (Benyshek et al., 1988); Henderson, 1988; Van Vleck, 1990a,b; Cantet et al., 1991a,b) for a vector of record $\boldsymbol{y}$ can be written as

$$\boldsymbol{y} = \mathbf{X}\beta + \boldsymbol{Z}_o\boldsymbol{a}_o + \boldsymbol{Z}_m\boldsymbol{a}_m + \boldsymbol{E}_m\boldsymbol{e}_m + \epsilon \quad [1]$$

where $\boldsymbol{y}$ is an n × 1 vector of records and $\mathbf{X}$ is an n × p incidence matrix that relates data to the unknown vector of location parameters β. The incidence matrices $\boldsymbol{Z}_o$, $\boldsymbol{Z}_m$, and $\boldsymbol{E}_m$ relate the unknown random vectors of direct BV ($\boldsymbol{a}_o$), maternal BV ($\boldsymbol{a}_m$), and maternal environmental effects ($\boldsymbol{e}_m$), respectively, to $\boldsymbol{y}$. The unknown vector ϵ contains the random residuals due to environmental effects peculiar to individual records.

There are a individuals with a direct BV in $\boldsymbol{a}_o$ and a maternal BV in $\boldsymbol{a}_m$. These a × 1 random vectors can be written as

$$\begin{bmatrix} \boldsymbol{a}_o \\ \boldsymbol{a}_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}_o\boldsymbol{g}_o \\ \boldsymbol{Q}_m\boldsymbol{g}_m \end{bmatrix} + \begin{bmatrix} \boldsymbol{a}_o^* \\ \boldsymbol{a}_m^* \end{bmatrix}$$

where $\boldsymbol{Q}_o$(a × n_{go}) and $\boldsymbol{Q}_m$(a × n_{gm}) are known incidence matrices relating BV to their respective means: the n_{go} elements of $\boldsymbol{g}_o$ and the n_{gm} elements of $\boldsymbol{g}_m$. The matrices $\boldsymbol{Q}_o$ and $\boldsymbol{Q}_m$ may be equal, as in Van Vleck (1990a,b), or not, as in Cantet (1990). The latter implies that different criteria are used to group animals for direct and for maternal effects. Differential grouping may be used in those situations in which the additive genetic trend for direct effects is previously observed to be different from the additive trend for maternal effects, as in Benyshek et al. (1988) or Cantet (1990).

Direct BV for individuals without records and maternal BV for non-dams are predicted by means of the correlated structure of $[\boldsymbol{a}_o' | \boldsymbol{a}_m']'$, using the method of Henderson (1977) when including random variables not in the model for records. The random vectors $\boldsymbol{a}_o^*$ and $\boldsymbol{a}_m^*$ are such that

$$\mathbf{E}\begin{bmatrix} \boldsymbol{a}_o^* \\ \boldsymbol{a}_m^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\mathbf{Var}\begin{bmatrix} \boldsymbol{a}_o^* \\ \boldsymbol{a}_m^* \end{bmatrix} = \begin{bmatrix} \sigma^2_{Ao} & \sigma_{AoAm} \\ \sigma_{AoAm} & \sigma^2_{Am} \end{bmatrix} \otimes \boldsymbol{A} \quad [3]$$

The a × a positive-definite matrix $\boldsymbol{A}$ contains the additive relationships among animals. Dispersion parameters σ^2_{Ao}, σ^2_{Am}, and σ_{AoAm} are the variances of additive direct and additive maternal effects and their covariance, respectively. Let $\boldsymbol{G}_0$ be the 2 × 2 matrix in [3] that contains these parameters. Hence, we have

$$\mathbf{E}\begin{bmatrix} \boldsymbol{a}_o \\ \boldsymbol{a}_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}_o\boldsymbol{g}_o \\ \boldsymbol{Q}_m\boldsymbol{g}_m \end{bmatrix}$$

$$\mathbf{Var}\begin{bmatrix} \boldsymbol{a}_o \\ \boldsymbol{a}_m \end{bmatrix} = \boldsymbol{G0} \otimes \boldsymbol{A} \quad [4]$$

The vectors $\boldsymbol{e}_m$ and ϵ are assumed to have zero expectations, to be independent of each other, and to be independent of the vectors of BV. The variance-covariance matrix of $\boldsymbol{e}_m$ is $\boldsymbol{I}_{n_d}\sigma^2_{Em}$. The scalar n_d is equal to the number of known dams. The matrix $\boldsymbol{e}_m$ is of order n × n_d. The variance of ϵ is $\boldsymbol{R}$, a positive-definite diagonal matrix with non-zero elements equal to σ^2_{Eo}, for those individuals with a record in $\boldsymbol{y}$ and with identified dam, and to $\sigma^2_{Em} + \sigma^2_{Eo}$ for individuals with a record in $\boldsymbol{y}$ but a missing dam.

Using all these specifications, the expected value of $\boldsymbol{y}$ in [1] is

$$\mathbf{E}(\boldsymbol{y}) = \boldsymbol{X\beta} + \boldsymbol{Z}_o\boldsymbol{Q}_o\ \boldsymbol{g}_o + \boldsymbol{Z}_m\boldsymbol{Q}_m\boldsymbol{g}_m \quad [5]$$

and the variance of $\boldsymbol{y}$ is

$$\mathbf{Var}(\boldsymbol{y}) = \boldsymbol{Z}_o\boldsymbol{A}\boldsymbol{Z}_o'\sigma^2_{Ao} + \boldsymbol{Z}_m\boldsymbol{A}\boldsymbol{Z}_m'\sigma^2_{Am} + [\boldsymbol{Z}_o\boldsymbol{A}\boldsymbol{Z}_m' + \boldsymbol{Z}_m\boldsymbol{A}\boldsymbol{Z}_o']\sigma_{AoAm} + \boldsymbol{E}_m\boldsymbol{E}_m'\sigma^2_{Em} + \boldsymbol{R} \quad [6]$$

Specifying Var(y) for Individuals with Missing Dams

A problem when writing Var($\boldsymbol{y}$) is that the additive part of the variances, or the covariances between individuals, may be mis-specified when information on some of their dams is missing. Lack of dam identification may lead to expressions for Var($\boldsymbol{y}$) that do not correspond with the assumptions of the model. The genetic part of the record of individual i with dam j in model [1] is $a_{oi} + a_{mj}$. The additive covariance between individuals i and k with dams j and l (Willham, 1963) is

$$\mathrm{cov}(a_{oi} + a_{mj}, a_{ok} + a_{ml}) = \boldsymbol{A}_{ik}\sigma^2_{Ao} + [\boldsymbol{A}_{il} + \boldsymbol{A}_{kj}]\sigma_{AoAm} + \boldsymbol{A}_{jl}\sigma^2_{Am} \quad [7]$$

where $\boldsymbol{A}_{ij}$ is element i,j of $\boldsymbol{A}$. Expression [7] shows that the additive covariance between relatives when maternal effects are present is a function of the additive relationships $\boldsymbol{A}_{ik}$, $\boldsymbol{A}_{il}$, $\boldsymbol{A}_{jk}$, and $\boldsymbol{A}_{jl}$. Conversely, the additive covariance between relatives when maternal effects are absent is a function of only one additive relationship: $\boldsymbol{A}_{ik}$. When dam identification is missing, the additive relationships between the dam of the individual and other individuals are computed as 0, zeroing out (partially or completely) all associated parameters from the expectation. As an example, consider individuals S, D, and O, the last two with records. Parents of S and D are unidentified. S is a male, D a female, and both are the parents of O. Ordering the vectors of records as $[y_D|y_O]'$, of BV for direct effects as $[a_{oS}|a_{oD}|a_{oO}]'$ and of BV for maternal effects as $[a_{mS}|a_{mD}|a_{mO}]'$, the matrices $\boldsymbol{Z}_o$, $\boldsymbol{Z}_m$ and $\boldsymbol{A}$ are

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

$$\boldsymbol{A} = \begin{bmatrix} 1.00 & 0 & .50 \\ 0 & 1.00 & .50 \\ .50 & .50 & 1.00 \end{bmatrix}$$

The matrices of Var($\boldsymbol{y}$) associated with σ^2_{Ao}, σ^2_{Am}, and σ_{AoAm}, respectively, are

$$\boldsymbol{Z}_o\boldsymbol{A}\boldsymbol{Z}_o' = \begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix} \qquad \boldsymbol{Z}_m\boldsymbol{A}\boldsymbol{Z}_m' = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\boldsymbol{Z}_o\boldsymbol{A}\boldsymbol{Z}_m' + \boldsymbol{Z}_m\boldsymbol{A}\boldsymbol{Z}_o' + \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

Using [7] to check these matrices indicates that only σ^2_{Ao} is properly accounted because σ^2_{Am} is lacking from Var($\boldsymbol{y}_D$) and cov ($\boldsymbol{y}_D$,$\boldsymbol{y}_O$). Also, σ_{AoAm} is lacking from Var($\boldsymbol{y}_D$), whereas its coefficient in cov($\boldsymbol{y}_D$,$\boldsymbol{y}_O$) is 1 instead of the correct 1.25 (Willham, 1963). One way of solving the problem is to write the remaining fractions of the coefficients of σ^2_{Am} and σ_{AoAm} that are unaccounted due to lack of dam identification into two matrices that are added to $\boldsymbol{R}$, for example. However, this complicates computing the MME due to the need of

inverting a nondiagonal $\boldsymbol{R}$. Another possibility is to enlarge the vectors $\boldsymbol{a}_o$ and $\boldsymbol{a}_m$ with the BV of the missing or "phantom" dams (Westell et al., 1988) for all those individuals with records in $\boldsymbol{y}$ and unidentified dams as described by Van Vleck (1990a,b). "Phantom" dams are assumed to be unrelated and to have one single progeny in either $\boldsymbol{a}_o$ or $\boldsymbol{a}_m$ (Quaas, 1988; Westell et al., 1988). In the example the "phantom" dam of D (N, say) is included in $\boldsymbol{a}_o$ and $\boldsymbol{a}_m$ such that, if animals are ordered N, S, D, and O, we have

$$Z_o = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad Z_m = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 0 & .5 & .25 \\ 0 & 1 & 0 & .5 \\ .5 & 0 & 1 & .5 \\ .25 & .5 & .5 & 1 \end{bmatrix}$$

With this enlargement the correct matrices for σ^2_{Am} and σ_{AoAm} in Var($\boldsymbol{y}$) are obtained:

$$Z_m A Z_m' = \begin{bmatrix} 1 & .5 \\ .5 & 1 \end{bmatrix}$$

$$Z_o A Z_m' + Z_m A Z_o' = \begin{bmatrix} 1 & 1.25 \\ 1.25 & 1 \end{bmatrix}$$

It should be pointed out that only BV of "phantom" dams of individuals with records in y have to be incorporated in $\boldsymbol{a}_o$ and $\boldsymbol{a}_m$. Note that a missing dam for S does not help writing Var($\boldsymbol{y}$) properly, in the example, and increases the order of the vectors of BV unnecessarily.

For individuals with unidentified dams, σ^2_{Em} should also be added to the diagonal elements of Var($\boldsymbol{y}$) (Henderson, 1988). To force the variance of any record to include σ^2_{Em}, residuals in [1] can be written as

$$\begin{aligned} \epsilon &= M e_m^M + e_o \\ E(\epsilon) &= 0 \\ Var(\epsilon) &= R = MM'\sigma^2_{Em} + I_n\sigma^2_{Eo} \end{aligned} \qquad [8]$$

Let n_h be the number of "phantom" dams included in $\boldsymbol{a}_o$ or $\boldsymbol{a}_m$. Then, matrix $\boldsymbol{M}$ of order $n \times n_h$ associates records to the environmental deviations of "phantom" dams (e_m^M). Because every missing dam has only one progeny, off-diagonals and diagonal elements of MM' for individuals with a known dam are equal to zero, whereas diagonal elements for individuals with missing dams are equal to 1. Observe that $\boldsymbol{R}$ in [6] can be written such that $R = MM'\sigma^2_{Em} + I_n\sigma^2_{Eo}$.

Reduced Animal Model with Genetic Grouping

Let n_g be the number of groups for both direct and maternal effects. Then, the number of equations in the AM with genetic grouping [1] is $(p + n_g + 2(a + n_h) + n_d)$. A sizeable reduction in this number is possible by using a RAM with genetic grouping if the number of "phantom" dams that are to be included in $\boldsymbol{a}_o$ and $\boldsymbol{a}_m$ is small and the number of non-parents is large relative to the number of parents. Ordinary RAM requires partitioning the data vector $\boldsymbol{y}$ into n_P records of individuals with progeny ($\boldsymbol{y}_P$; parents) and n_N records of individuals without progeny ($\boldsymbol{y}_N$; non-parents) so that $\boldsymbol{y}' = [\boldsymbol{y}_P', \boldsymbol{y}_N']$. A conformable partition can be used in $\boldsymbol{X}$, $\boldsymbol{Z}_o$, $\boldsymbol{Z}_m$, $\boldsymbol{Q}_o$, $\boldsymbol{Q}_m$, $\boldsymbol{e}_m$, $\boldsymbol{M}$, $\boldsymbol{A}$, $\boldsymbol{a}_o$, $\boldsymbol{a}_m$, and ϵ. The incidence matrices for BV in [1] (after enlarging $\boldsymbol{a}_o$ and $\boldsymbol{a}_m$ with "phantom" dams for individuals with records) are

$$Z_o = \begin{bmatrix} \mathbf{0}_{n_p x n_h} & Z_o^P & \mathbf{0}_{n_p x n_N} \\ \mathbf{0}_{n_N x n_h} & \mathbf{0}_{n_N x n_P} & Z_o^N \end{bmatrix}$$

$$Z_m = \begin{bmatrix} M^P & Z_m^P & \mathbf{0}_{n_p x n_N} \\ M^N & Z_m^{NP} & \mathbf{0}_{n_N x n_N} \end{bmatrix} \qquad [9]$$

Matrix $\boldsymbol{Z}_m$ has columns of zeros except for those relating a record of a parent or a non-parent with the maternal BV of its identified or "phantom" dam. Note that $\boldsymbol{Z}_m$ contains zero submatrices associated with non-parents with no recorded progeny themselves. Submatrices $\boldsymbol{M}^P$ and Z_m^P relate records of parents to maternal BV of "phantom" and identified dams, respectively. Corresponding association between records of non-parents to maternal BV of "phantom" and identified dams is by means of $\boldsymbol{M}^N$ and Z_m^{NP}.

Appendix A presents the matrices involved in the partitions of $\boldsymbol{a}_o$, $\boldsymbol{a}_m$, $\boldsymbol{Q}_o$, $\boldsymbol{Q}_m$, and $\boldsymbol{A}$. To obtain equations for RAM with genetic grouping, the BV of "base" parents are ordered such that "phantom" sires precede "phantom" dams. Hence, $P_b = [P_{bS}^N | P_{bD}^N]$. It is assumed that each non-parent has only one record in $\boldsymbol{y}$ (i.e., $Z_o^N = I_{nN}$), a reasonable assumption for maternally influenced traits that are usually observed once during the lifetime of the individual. After all these specifications, the

following equivalent model (Henderson, 1985) to [1] can be written:

$$\begin{bmatrix} y_P \\ y_N \end{bmatrix} = \begin{bmatrix} X_P \\ X_N \end{bmatrix} \beta + \begin{bmatrix} 0 & Z_o^P \\ P_b^N & P^{NP} \end{bmatrix} \begin{bmatrix} Q_{bo} \\ Q_o^P \end{bmatrix} g_o$$

$$+ \begin{bmatrix} M^P & Z_m^P \\ M^N & Z_m^{NP} \end{bmatrix} \begin{bmatrix} Q_m^D \\ Q_m^P \end{bmatrix} g_m \begin{bmatrix} 0 & Z_o^P \\ P_{bD}^N & P^{NP} \end{bmatrix} \begin{bmatrix} a_{bo}^{D*} \\ a_o^{P*} \end{bmatrix}$$

$$+ \begin{bmatrix} M^P & Z_m^P \\ M^N & Z_m^{NP} \end{bmatrix} \begin{bmatrix} a_{bm}^{D*} \\ a_m^{P*} \end{bmatrix} + \begin{bmatrix} E_m^P \\ E_m^N \end{bmatrix} e_m$$

$$+ \begin{bmatrix} \epsilon^P \\ \epsilon^N + \phi_o^N + P_{bS}^N a_{bo}^{S*} \end{bmatrix} \quad [10]$$

where the matrix Q_m^D is such that $a_{bm}^D = Q_m^D g_m + a_{bm}^{D*}$. Also, $E(a_{bm}^{D*}) = 0$ so that $E(a_{bm}^D) = Q_m^D g_m$. Note that the model equation is expressed in terms of a_{bm}^{D*} rather than a_{bm}^D, the latter being the maternal BV of the "phantom" dams.

An important characteristic of [10] is that the variance-covariance structure of the residuals is diagonal and uncorrelated with other random vectors in the model:

$$\text{Var} \begin{bmatrix} \epsilon^P \\ \epsilon^N + \Phi_o^N + P_{bS}^N a_{bo}^{S*} \end{bmatrix} = \begin{bmatrix} R_P & 0 \\ 0 & R_N \end{bmatrix} =$$

$$\begin{bmatrix} M^P M^{P'} \sigma_{Em}^2 & 0 \\ + I_{n_P} \sigma_{Eo}^2 & \\ 0 & (P_{bS}^N P_{bS}^{N'} + \frac{1}{2} I_{n_N}) \sigma_{Ao}^2 + M^N M^{N'} \sigma_{Em}^2 + I_{n_N} \sigma_{Eo}^2 \end{bmatrix}$$

Appendix B presents a proof that [1] and [10] are equivalent models. The QP-transformation of Quaas and Pollak (1981) is used to obtain MME for model [10] that are amenable to calculation. Let the vector of QP-transformed solutions for [10] be:

$$\hat{\theta}' = [\hat{\beta}' | \hat{g}_o' | \hat{g}_m' | \hat{a}_{bo}^{D'} | \hat{a}_o^{P'} | \hat{a}_{bm}^{D'} | \hat{a}_m^{P'} | \hat{e}_m']$$

and define:

$$F_P = [X_P | 0_{n_P x n_g} | 0_{n_P x n_h} | Z_o^P | M^P | Z_m^P | E_m^P]$$

$$F_N = [X_N | 0_{n_N x n_g} | P_{bD}^N | P^{NP} | M_N | Z_m^{NP} | E_m^N]$$

Then, QP-transformed equations for model [10] are

$$[F_P' R_P^{-1} F_P + F_N' R_N^{-1} F_N + A^*] \hat{\theta} = F_P' R_P^{-1} y_P + F_N' R_N^{-1} y_N \quad [11]$$

where

$$A^* = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & Q^{RM'}(G_0^{-1} \otimes A_P^{-1}) Q^{RM} & -Q^{RM'}(G_0^{-1} \otimes A_P^{-1}) & 0 \\ 0 & -(G_0^{-1} \otimes A_P^{-1}) Q^{RM} & G_0^{-1} \otimes A_P^{-1} & 0 \\ 0 & 0 & 0 & I_{n_d}(\sigma_{Em}^2)^{-1} \end{bmatrix} \sigma_{Eo}^2$$

The matrix A_P contains the additive relationships among the "phantom" dams of animals with records and known parents. Also:

$$Q^{RM} = \begin{bmatrix} Q_{bo}^D & 0 \\ Q_o^P & 0 \\ 0 & Q_{bm}^D \\ 0 & Q_m^P \end{bmatrix}$$

In practice, the part of the system of equations [11] that involves F-matrices may be formed as a weighted (by a diagonal R) least squares matrix while reading the data file. The central block of A^* can be calculated by the rules described by Cantet (1990), which extend those of Westell et al. (1988) and Quaas (1988) when maternal effects are absent.

Backsolving for Non-Parents

The BV for direct and maternal effects for non-parents can be calculated by first solving Equations [11] and then, by replacing random variables in [A4] by their corresponding predictors.

The approach of Henderson (1985, 1988) will be used to calculate the Mendelian residuals for direct and maternal effects. Observe that

$$\text{BLUP}(\phi_o^N) = \hat{\phi}_o^N = \text{cov}(\phi_o^N, \epsilon^{N'}) \text{Var}(\epsilon^N)^{-1} \text{BLUP}(\epsilon^N)$$

where $\text{BLUP}(\epsilon^N) = \hat{\epsilon} = y^N - F^N \hat{\theta}$. Therefore, the direct Mendelian residuals for non-parents can be predicted by means of

$$\phi_o^N = \frac{1}{2} \sigma_{Ao}^2 \left(\left(P_{bS}^N P_{bS}^{N'} + \frac{1}{2} I_{n_N} \right) \sigma_{Ao}^2 + M^N M^{N'} \sigma_{Em}^2 + I_{n_N \sigma_{Eo}^2} \right)^{-1}$$

$$(y_N - F_N \hat{\theta}) \quad [12]$$

The fact that the matrix on the right of [12] is diagonal allows obtaining the direct Mendelian

residual for non-parent i using scalar operations as follows:

$$\hat{\phi}^N_{o_i} = \left[\frac{\frac{1}{2}\sigma^2_{Ao}}{(s_i + \frac{1}{2})\sigma^2_{Ao} + m_i\sigma^2_{Em} + \sigma^2_{Eo}} \right] (y_i - \boldsymbol{F}_{N_i}\hat{\boldsymbol{\theta}})$$

The scalar s_i is equal to ¼ if i has an unidentified sire and 0 otherwise. The scalar m_i is the diagonal element of $\mathbf{MM}'$ corresponding to non-parent i and is equal to 1 if i has a missing dam and equal to 0 if the dam is known. Vector $\boldsymbol{F}_{Ni}$ is row i of $\boldsymbol{F}_N$. After [12] is calculated the maternal Mendelian residuals for non-parents are obtained by

$$\text{BLUP}(\boldsymbol{\phi}^N_m) = \hat{\boldsymbol{\phi}}^N_m = \text{co,v}\boldsymbol{\phi}^N_m,\boldsymbol{\phi}^{N'}_o)\text{Var}(\boldsymbol{\phi}^N_o)^{-1}\text{BLUP}(\boldsymbol{\phi}^N_o) = \left(\frac{\sigma_{AoAm}}{\sigma^2_{Ao}} \right) \hat{\boldsymbol{\phi}}^N_o \quad [13]$$

Appendix C presents the derivation of the predictors of BV for non-parents, which are equal to

$$\begin{bmatrix} \hat{\boldsymbol{a}}^N_o \\ \hat{\boldsymbol{a}}^N_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}^{NS}_{bo} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{Q}^{NS}_{bm} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{g}}_o \\ \hat{\boldsymbol{g}}_m \end{bmatrix} + (\boldsymbol{I}_2 \otimes \boldsymbol{P}^{NP}_{bD}) \begin{bmatrix} \hat{\boldsymbol{a}}^{P*}_{bo} \\ \hat{\boldsymbol{a}}^{D*}_{bm} \end{bmatrix}$$

$$+ (\boldsymbol{I}_2 \otimes \boldsymbol{P}^{NP}) \begin{bmatrix} \hat{\boldsymbol{a}}^P_o \\ \hat{\boldsymbol{a}}^P_m \end{bmatrix} + 2 \begin{bmatrix} \hat{\phi}^N_o \\ \hat{\phi}^N_m \end{bmatrix} \quad [14]$$

Example

Consider a pedigree with 11 individuals, 5 of them are parents and 6 non-parents. Capital letters denote the individuals (all of them having records) and small letters the unknown or "phantom" parents. Numbers in parentheses indicate the groups to which the unknown parents belong for direct and maternal effects. The pedigree file is as follows:

Individual	Sire	Dam	Parentage
A	a(1,1)	b(1,2)	parent
B	A	e(2,1)	parent
C	f(2,1)	g(1,1)	parent
D	h(2,2)	i(2,2)	parent
E	c(1,2)	d(2,1)	parent
F	B	E	non-parent
G	B	C	non-parent
H	B	C	non-parent
I	j(2,2)	C	non-parent
J	A	D	non-parent
K	k(1,2)	l(2,1)	non-parent

Let the records for A to K be 228, 264, 213, 209, 210, 190, 210, 260, 215, 230, and 191, respectively. Sex is the only fixed effect in the model; A, B, F, H, and J are males and C, D, E, G, I, and K are females.

We now show the different matrices involved in the transformed RAM with genetic grouping Equations [11]. The matrix $\boldsymbol{F}_P$ that relates records of parents to the vector of solutions is as follows:

$$\boldsymbol{F}_P = [\boldsymbol{X}_P|\mathbf{0}_{5\times4}|\mathbf{0}_{5\times6}|\boldsymbol{I}_5|\boldsymbol{M}^P|\mathbf{0}_{5\times5}|\mathbf{0}_{5\times3}]$$

The matrix M^P of order 5 (number of parents) by 6 (number of "phantom" dams) is equal to

$$\boldsymbol{M}^P = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array} \begin{array}{c} \begin{array}{cccccc} b & d & e & g & i & l \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

The matrix $\boldsymbol{F}_N$ relating records of non-parents to the vector of solutions is equal to

$$\boldsymbol{F}_N = [\boldsymbol{X}_N|\mathbf{0}_{6\times4}|\boldsymbol{P}^N_{bD}|\boldsymbol{P}^{NP}|\boldsymbol{M}^N|\boldsymbol{Z}^{NP}_m|\boldsymbol{E}^N_m]$$

Because there is only one non-parent (K) with a "phantom" dam (l), the matrices $\boldsymbol{P}^N_{bD}$ and $\boldsymbol{M}^N$ are of order 6 (number of "phantom" dams) by 6 (number of non-parents) with all elements equal to zero except for the (6,6) which is equal to .5 for $\boldsymbol{P}^N_{bD}$, and equal to 1 for $\boldsymbol{M}^N$.

The matrices $\boldsymbol{P}^{NP}$, which relates BV of non-parents to the parental ones, and $\boldsymbol{Z}^{NP}_m$, which relates records on non-parents to maternal BV of parents, are of the same order (6 × 5) and are equal to

$$\boldsymbol{P}^{NP} = \begin{array}{c} \\ F \\ G \\ H \\ I \\ J \\ K \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} .0 & .5 & .0 & .0 & .5 \\ .0 & .5 & .5 & .0 & .0 \\ .0 & .5 & .5 & .0 & .0 \\ .0 & .0 & .5 & .0 & .0 \\ .5 & .0 & .0 & .5 & .0 \\ .0 & .0 & .0 & .0 & .0 \end{bmatrix} \end{array}$$

$$\boldsymbol{Z}^{NP}_m = \begin{array}{c} \\ F \\ G \\ H \\ I \\ J \\ K \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Observe that Z_m^{NP} has columns equal to zero for the sires (A and B) and equal to two times the columns of P^{NP} corresponding to dams (C, D, and E).

The matrix of maternal environmental effects for records of non-parents is as follows:

$$E_m^N = \begin{array}{c} \\ F \\ G \\ H \\ I \\ J \\ K \end{array} \begin{array}{c} \begin{array}{ccc} C & D & E \end{array} \\ \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

The central block of A^* in [11] is a function of A_P^{-1} and of the following matrices:

$$Q_{bo}^D = \begin{array}{c} \\ b \\ d \\ e \\ g \\ i \\ l \end{array} \begin{array}{c} \begin{array}{cc} g_{o_1} & g_{o_2} \end{array} \\ \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \end{array} \qquad Q_{bm}^D = \begin{array}{c} \\ b \\ d \\ e \\ g \\ i \\ l \end{array} \begin{array}{c} \begin{array}{cc} g_{m_1} & g_{m_2} \end{array} \\ \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \end{array}$$

$$Q_o = \begin{bmatrix} Q_{Po} \\ -- \\ Q_{No} \end{bmatrix} = \begin{array}{c} A \\ B \\ C \\ D \\ E \\ - \\ F \\ G \\ H \\ I \\ J \\ K \end{array} \begin{array}{c} \begin{array}{cc} g_{o_1} & g_{o_2} \end{array} \\ \begin{bmatrix} 1.00 & .00 \\ .50 & .50 \\ .50 & .50 \\ .00 & 1.00 \\ .50 & .50 \\ -- & -- \\ .50 & .50 \\ .50 & .50 \\ .50 & .50 \\ .25 & .75 \\ .50 & .50 \\ .50 & .50 \end{bmatrix} \end{array}$$

$$Q_m = \begin{bmatrix} Q_{Po} \\ Q_{Pm} \end{bmatrix} = \begin{array}{c} A \\ B \\ C \\ D \\ E \\ - \\ F \\ G \\ H \\ I \\ J \\ K \end{array} \begin{array}{c} \begin{array}{cc} g_{m_1} & g_{m_2} \end{array} \\ \begin{bmatrix} .500 & .500 \\ .750 & .250 \\ 1.000 & .000 \\ .000 & 1.000 \\ .500 & .500 \\ -- & -- \\ .625 & .375 \\ .875 & .125 \\ .875 & .125 \\ .500 & .500 \\ .250 & .750 \\ .500 & .500 \end{bmatrix} \end{array}$$

The (co)variance components are $\sigma^2_{Ao} = .5$, $\sigma^2_{Am} = .30$, $\sigma_{AoAm} = -.20$, $\sigma^2_{Em} = .25$, and $\sigma^2_{Eo} = 1.00$. The residual variances of parents and non-parents are $R_P = I_5$ 1.25 and R_N = Diag(1.25, 1.25, 1.25, 1.25, 1.25, 1.50), respectively.

Restrictions imposed to solve the equations were to set to zero group 2 of direct effects and group 2 of maternal effects. The resulting solutions were males = 234.578, females = 197.1; group 1 of direct effects = −18.1232, and group 1 of maternal effects = 20.176; BV for direct effects: "phantom" dams: b = −17.5117, d = −.516515, e = −1.93473, g = −17.8636, i = .346237, l = −.465241, parents A to E = −15.1093, −8.18981, −7.91753, 2.59371, −9.1491; BV for maternal effects: "phantom" dams: b = 1.06884, d = 18.1099, e = 23.4435, g = 21.2143, i = 1.38945, l = 18.315, parents A to E = 9.53915, 16.3584, 21.3153, −.0819643, 4.5376; maternal environmental effects: C = 1.55437, D = .293636, E = −6.74102. Mendelian residuals of non-parents F to K for direct effects (Equations [12]) are as follows: −6.741046, −.38311, 2.1211999, −.183716, .2936405, −2.326193; and for maternal effects (Equations [13]): 2.696406, .1532442, −.848481, .0734879, −.117454, and .093482. Finally, BV for direct effects of non-parents F to K predicted by means of [14] are −15.4105, −8.43678, −5.93247, −4.23434, −5.96416, and −12.7835, respectively. Corresponding maternal BV are 13.1444, 18.9901, 17.9884, 10.7679, 4.61114, and 10.5532.

Discussion

The equations for RAM with genetic grouping with maternal effects obtained here, as well as the ones of Van Vleck (1990b), require the inclusion of BV for "phantom" dams in a_o and a_m. As a consequence, the number of animals to be evaluated increases and so the computational burden. Including the BV of unidentified dams is a way of avoiding mis-specification of σ^2_{Am} and σ_{AoAm}. Westell et al. (1988) absorbed the BV of missing dams for the situation in which maternal effects are absent. In this case the absorption only takes place in the A^* part of the system. However, when maternal effects are fitted in the model, absorption of the BV of phantom dams involves submatrices of both A^* and $F_P' R_P^{-1} F_P$ of [11]. At this time it is not clear whether the absorption can be computed in a feasible way and additional research is needed in this direction.

If groups for direct and for maternal effects are to be assigned under the same criterion, the results will be the same as in Van Vleck (1990b), if the same constraints are imposed to solve the equations. From the point of view of estimation in linear models, assigning the same groups for direct and maternal effects to missing parents

produces a replicated set of columns in the estimation space (Zyskind, 1969). This, in turn, makes solutions of groups for direct and maternal effects difficult to interpret, as observed by Van Vleck (1990a,b). More discussion on this topic can be found in Cantet (1990). Conversely, differential assignment of groups for direct and maternal effects seems to be useful when different genetic trends for both types of effects are previously observed. For example, Benyshek et al. (1988) and Cantet (1990) found a positive genetic trend for additive direct effects, whereas the trend for additive maternal effects was zero. In this situation, only one group would be needed for maternal effects.

Implications

The model with different assignment of genetic groups for direct and maternal effects is appropriate for estimating direct and maternal breeding values when the additive genetic trend for direct effects has been previously observed to be different from the additive genetic trend for maternal effects. The reduced animal model presented here, that incorporates different direct and maternal genetic groups into the mixed-model equations, allows computing breeding values of parents and non-parents with currently used algorithms. Inclusion of "phantom" dams in $\boldsymbol{a}_o$ and $\boldsymbol{a}_m$ avoids misspecification of maternal additive variance and additive covariance between direct and maternal effects, when dam information is missing. Specification of these covariance components should be carefully considered, especially if they are to be estimated from the data.

Literature Cited

Benyshek, L. L., M. H. Johnson, D. E. Little, J. K. Bertrand, and L. A. Kriese. 1988. Applications of an animal model in the United States beef cattle industry. J. Dairy Sci. 71(Suppl. 2): 35.

Cantet, R.J.C. 1990. Estimation and prediction from mixed linear models for maternal effects. Ph.D. Thesis. Univ. of Illinois, Urbana.

Henderson, C. R. 1973. Sire evaluation and genetic trends. Proc. Anim. Breed. Genet. Symp. in Honor of Dr. J. L. Lush, ADSA and ASAS, Champaign, IL.

Henderson, C. R. 1976. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics 32:69.

Henderson, C. R. 1977. Best linear unbiased prediction of breeding values not in the model for records. J. Dairy Sci. 60:783.

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

Henderson, C. R. 1988. Theoretical basis and computational methods for a number of different animal models. J. Dairy Sci. 71(Suppl. 2):1.

Henderson, C. R., and R. L. Quaas. 1976. Multiple trait evaluation using relatives' records. J. Anim. Sci. 43:1188.

Quaas, R. L. 1988. Additive genetic model with groups and relationships. J. Dairy Sci. 71:1338.

Quaas, R. L., and E. J. Pollak. 1980. Mixed model methodology for farm and ranch beef cattle testing programs. J. Anim. Sci. 51:1277.

Quaas, R. L., and E. J. Pollak. 1981. Modified equations for sire models with groups. J. Dairy Sci. 64:1868.

Robinson, G. K. 1986. Group effects and computing strategies for models for estimating breeding values. J. Dairy Sci. 69: 3106.

Taylor, J. F., and M. A. Tomaszewski. 1989. Inverse numerator relationship matrix approximation. J. Dairy Sci. 72:664.

Van Vleck, L. D. 1990a. Breeding value prediction with maternal genetic groups. J. Anim. Sci. 68:3998.

Van Vleck, L. D. 1990b. Absorption of equations for non-parents for an animal model with maternal effects and genetic groups. J. Anim. Sci. 68:4014.

Westell, R. A., R. L. Quaas, and L. D. Van Vleck. 1988. Genetic groups in animal model. J. Dairy Sci. 71:1310.

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

Zyskind, G. 1969. Parametric augmentations and error structures under which certain simple least-squares and analysis of variance procedures are also best. Ann. Math. Stat. 40:1353.

APPENDIX A: Expressions for the Partitions of a_o, a_m, Q_o, Q_m, and A

The BV of any individual can be expressed as a linear function of the BV of individuals in previous generations plus a residual term due to Mendelian sampling (Quaas, 1988). Reference to three conceptual "generations" is needed when obtaining a RAM with genetic grouping, and these make up the "base" individuals plus missing parents from further generations, the parents and the non-parents, with vectors of direct BV equal to $\boldsymbol{a}_{bo}$, $\boldsymbol{a}_o^P$, and $\boldsymbol{a}_o^N$, respectively. Corresponding BV for maternal effects are $\boldsymbol{a}_{bm}$, $\boldsymbol{a}_m^P$, and $\boldsymbol{a}_m^N$. Put $\boldsymbol{a}_o = [\boldsymbol{a}_o^{P'}|\boldsymbol{a}_o^{N'}]'$, and $\boldsymbol{a}_m = [\boldsymbol{a}_m^{P'}|\boldsymbol{a}_m^{N'}]'$. It is assumed that every individual in the "base" population has only one progeny and there is no inbreeding, as in Quaas (1988). Therefore, the direct and maternal BV can be written as:

$$\begin{bmatrix} \boldsymbol{a}_o \\ \boldsymbol{a}_m \end{bmatrix} = [\boldsymbol{I}_2 \otimes [\boldsymbol{P}_b|\boldsymbol{P}]] \begin{bmatrix} \boldsymbol{a}_{bo} \\ \boldsymbol{a}_o \\ \boldsymbol{a}_{bm} \\ \boldsymbol{a}_m \end{bmatrix} + \begin{bmatrix} \boldsymbol{\phi}_o \\ \boldsymbol{\phi}_m \end{bmatrix} \quad [A1]$$

where $\boldsymbol{P}_b$ is the matrix that relates individuals in the "base" to animals in $\boldsymbol{a}_o$ (or in $\boldsymbol{a}_m$), $\boldsymbol{P}$ is the matrix that relates individuals in $\boldsymbol{a}_o$ (or in $\boldsymbol{a}_m$), and $\boldsymbol{\phi}_o$ and $\boldsymbol{\phi}_m$ are vectors of individuals' deviations with respect to the mid-parental BV for direct and

maternal effects, respectively, caused by Mendelian sampling of genes (Quaas, 1988). After rearranging and solving for $[a_o'|a_m']'$, gives

$$\begin{bmatrix} a_o \\ a_m \end{bmatrix} = [I_2 \otimes (I_a - P)^{-1} P_b] \begin{bmatrix} a_{bo} \\ a_{bm} \end{bmatrix} + \begin{bmatrix} \phi_o \\ \phi_m \end{bmatrix} \quad [A2]$$

The non-symmetric matrix $[P_b|P]$ that describes the gene flow can be partitioned as

$$[P_b|P] = \left[\begin{array}{cc|cc} P_b^P & 0 & P^{PP} & 0 \\ 0 & P_b^N & P^{NP} & 0 \end{array}\right] \quad [A3]$$

where P_b^P and P_b^N relate the BV of parents and non-parents, respectively, to "base" individuals. The submatrix P^{PP} describes the relationship between the BV of parents and P^{NP} describes the relationship between the BV of parents and non-parents. The submatrices on the right of P are zero because there is no transmission of genes from non-parents to parents or among non-parents. If all non-parents have parents in a_o^P (or in a_m^P), P_b^N is a null matrix. Using these submatrices in [A1], the vector of non-parental BV for direct and maternal effects can be written as

$$\begin{bmatrix} a_o^N \\ a_m^N \end{bmatrix} = (I_2 \otimes P_b^N) \begin{bmatrix} a_{bo}^N \\ a_{bm}^N \end{bmatrix} + (I_2 \otimes P^{NP}) \begin{bmatrix} a_o^P \\ a_m^P \end{bmatrix} + \begin{bmatrix} \phi_o^N \\ \phi_m^N \end{bmatrix} \quad [A4]$$

The vector $[\phi_o^{N'}|\phi_m^{N'}]'$ contains the Mendelian segregation residuals of non-parents for direct and maternal effects.

The a × a matrix $(I_a - P)$ is

$$(I_a - P) = \begin{bmatrix} (I_{n_P} - P^{PP}) & 0 \\ -P^{NP} & I_{n_N} \end{bmatrix} \quad [A5]$$

It can be verified that the inverse of $(I_a - P)$ is equal to

$$(I_a - P)^{-1} = \begin{bmatrix} (I - P^{PP})^{-1} & 0 \\ P^{NP}(I - P^{PP})^{-1} & I \end{bmatrix} \quad [A6]$$

For the procedure of Westell et al. (1988) and Robinson (1986) in absence of maternal effects, Quaas (1988) showed that Q_o of can be represented such that $Q_o = (I_a - P)^{-1}P_b\, Q_{bo}$. The matrix $Q_{bo} = [Q_b^{P'}, Q_b^{N'}]'$ relates "base" individuals plus missing parents to their population means. The i,j element of Q_o is the expected fraction of the i^{th} animal's genes deriving from the j^{th} population. To obtain a suitable expression for RAM with genetic grouping requires that Q_o be partitioned into a submatrix associated with parents (Q_o^P) and another submatrix associated with non-parents (Q_o^N) such that the latter is a linear function of the former. A closer inspection of Q_o reveals that

$$Q_o = \begin{bmatrix} Q_o^P \\ Q_o^N \end{bmatrix} = \begin{bmatrix} (I_{n_P} - P^{PP})^{-1} & 0 \\ P^{NP}(I_{n_P} - P^{PP})^{-1} & I_{n_N} \end{bmatrix} \begin{bmatrix} P_b^P Q_{bo}^P \\ P_b^N Q_{bo}^N \end{bmatrix}$$

$$= \begin{bmatrix} (I_{n_p} - P^{PP})^{-1} P_b^P Q_{bo}^P \\ P^{NP}(I_{n_p} - P^{PP})^{-1} P_b^P Q_{bo}^P + P_b^N Q_{bo}^N \end{bmatrix}$$

$$= \begin{bmatrix} Q_o^P \\ P^{NP} Q_o^P + P_b^N Q_{bo}^N \end{bmatrix} \quad [A7]$$

The factor $P_b^N Q_b^N$ in Q_o^N accounts for those non-parents with one or two unknown parents to be associated with direct genetic groups.

By a similar reasoning Q_m can be partitioned into Q_m^P and Q_m^N such that

$$Q_m = \begin{bmatrix} Q_m^P \\ P^{NP} Q_m^P + P_b^N Q_{bm}^N \end{bmatrix} \quad [A8]$$

Associated with the partition of a in parental and non-parental BV, there is a corresponding partition of A in

$$A = \begin{bmatrix} A_{PP} & A_{PN} \\ A_{NP} & A_{NN} \end{bmatrix}$$

In the absence of inbreeding, Quaas (1988) showed that

$$A = (I_a - P)^{-1}(P_b P_b' + .5 I_{n_b})(I_a - P')^{-1}$$

Note that $P_b P_b' = \text{Diag}\{.25 m_i\}$, for $m_i = 0,1,2$ = the number of "base" parents of the i^{th} individual. This is due to P_b having at most two .5 terms in any row, the remaining elements in the row being equal to zero. The number of missing parents is n_b. Letting $D = P_b P_b' + .5 I_{n_b} = \text{Diag}\{.25 m_i + .5\}$, we have that

$$A = (I_a - P)^{-1}D(I_a - P')^{-1}$$

$$= \begin{bmatrix} (I_{n_P} - P^{PP})^{-1} & 0 \\ P^{NP}(I_{n_P} - P^{PP})^{-1} & I_{n_N} \end{bmatrix} \begin{bmatrix} D_{PP} & 0 \\ 0 & D_{NN} \end{bmatrix}$$

$$\begin{bmatrix} (I_{n_P} - P^{PP'})^{-1} & (I_{n_P} - P^{PP'})^{-1} P^{NP'} \\ 0 & I_{n_N} \end{bmatrix}$$

and, on noting that $A_{PP} = (I - P^{PP})^{-1} D_{PP}(I - P^{PP'})^{-1}$, A is:

$$\begin{bmatrix} A_{PP} & A_{PP}P^{NP'} \\ P^{NP}A_{PP} \; P^{NP} & A_{PP}P^{NP'} + D_{NN} \end{bmatrix} \quad [A9]$$

Therefore, the relationship matrix among non-parents A_{NN} can be expressed as a function of the relationship matrix among parents (A_{PP}), the transition matrix P^{NP} and a diagonal matrix ($P_b^N P_b^{N'} + .5I$) associated with the relationships between base individuals and non-parents plus the Mendelian residuals of non-parents.

Expressions [A1], [A2], [A4], [A7], and [A8] expands the work of Quaas (1988) to the situation in which there are maternal effects. Formulas [A3], [A5], [A6], and [A9], also based on Quaas (1988), provide an insight of the relationships between the elements of the additive relationship matrix A (and its inverse) in different generations.

APPENDIX B: Proof That [1] and [10] are Equivalent Models

Two linear models are equivalent if their expected values and variance-covariance matrices are equal (Henderson, 1985). First we show that E(y) under [1] and [10] is the same. Using the definitions of Z_o and Z_m in [9] and disregarding the random vectors of environmental deviations which have zero expectations, the E(y) in [1] is

$$E\begin{bmatrix} y_P \\ y_N \end{bmatrix} = E\left[\begin{bmatrix} X_P \\ X_N \end{bmatrix}\beta + \begin{bmatrix} 0 & Z_o^P & 0 \\ 0 & 0 & I_{n_N} \end{bmatrix}\begin{bmatrix} a_o^D \\ a_o^P \\ a_o^N \end{bmatrix} + \begin{bmatrix} M_P & Z_m^P & 0 \\ M^N & Z_m^{NP} & 0 \end{bmatrix}\begin{bmatrix} a_m^D \\ a_m^P \\ a_m^N \end{bmatrix}\right]$$

$$= \begin{bmatrix} X_P \\ X_N \end{bmatrix}\beta + \begin{bmatrix} Z_o^P E(a_o^P) \\ E(a_o^N) \end{bmatrix} + \begin{bmatrix} M^P & Z_m^P \\ M^N & Z_m^{NP} \end{bmatrix} E\begin{bmatrix} a_m^D \\ a_m^P \end{bmatrix}$$

$$= \begin{bmatrix} X_P \\ X_N \end{bmatrix}\beta + \begin{bmatrix} Z_o^P Q_o^P \\ Q_o^N \end{bmatrix} g_o + \begin{bmatrix} M^P & Z_m^P \\ M^N & Z_m^{NP} \end{bmatrix}\begin{bmatrix} Q_m^D \\ Q_m^P \end{bmatrix} g_m \quad [B1]$$

Observing that all random vectors on the right of model [10] have zero means, the E(y) is

$$E\begin{bmatrix} y_P \\ y_N \end{bmatrix} = E\left[\begin{bmatrix} X_P \\ X_N \end{bmatrix}\beta + \begin{bmatrix} 0 & Z_o^P \\ P_b^N & P^{NP} \end{bmatrix}\begin{bmatrix} Q_{bo} \\ Q_o^P \end{bmatrix} g_o + \begin{bmatrix} M^P & Z_m^P \\ M^N & Z_m^{NP} \end{bmatrix}\begin{bmatrix} Q_m^D \\ Q_m^P \end{bmatrix} g_m\right] \quad [B2]$$

$$Var\begin{bmatrix} y_P \\ y_N \end{bmatrix} = Var\left[\begin{bmatrix} 0 & Z_o^P & 0 \\ 0 & 0 & I_{n_N} \end{bmatrix}\begin{bmatrix} a_o^D \\ a_o^P \\ a_o^N \end{bmatrix} + \begin{bmatrix} M^P & Z_m^P & 0 \\ M^N & Z_m^{NP} & 0 \end{bmatrix}\begin{bmatrix} a_m^D \\ a_m^P \\ a_m^N \end{bmatrix} + \begin{bmatrix} E_m^P \\ E_m^N \end{bmatrix} e_m + \epsilon\right]$$

$$= Var\left[\begin{bmatrix} Z_o^P & 0 \\ 0 & I_{n_N} \end{bmatrix}\begin{bmatrix} a_o^P \\ a_o^N \end{bmatrix} + \begin{bmatrix} M^P & Z_m^P \\ M^N & Z_m^{NP} \end{bmatrix}\begin{bmatrix} a_m^D \\ a_m^P \end{bmatrix}\right] + E_m E_m' \sigma_{Em}^2 + R$$

$$= Var\left[\begin{bmatrix} Z_o^P & 0 \\ 0 & In_N \end{bmatrix}\begin{bmatrix} a_o^P \\ a_o^N \end{bmatrix} + \begin{bmatrix} M^P & Z_m^{NP} \\ M^N & \end{bmatrix}\begin{bmatrix} a_m^D \\ a_m^P \end{bmatrix}\right] + (E_m E_m' + MM')\sigma_{Em}^2 + I_{n_N}\sigma_{Eo}^2$$

$$= \text{Var}\left[\begin{bmatrix} Z_o^P & \mathbf{0} \\ \mathbf{0} & I_{n_N} \end{bmatrix}\begin{bmatrix} a_o^{P*} \\ a_o^{N*} \end{bmatrix} + \begin{bmatrix} M^P & Z_m^P \\ M^N & Z_m^{NP} \end{bmatrix}\begin{bmatrix} a_m^{D*} \\ a_m^{P*} \end{bmatrix}\right] + (E_m E_m' + MM')\sigma^2_{Em} + I_{n_N}\sigma^2_{Eo} \qquad [B3]$$

The last equality holding after [2].

The variance of [10] is

$$\text{Var}(y) = \text{Var}\left[\begin{bmatrix} \mathbf{0} & Z_o^P \\ P_{bD}^N & P^{NP} \end{bmatrix}\begin{bmatrix} a_{bo}^{D*} \\ a_o^{P*} \end{bmatrix} + \begin{bmatrix} M^P & Z_m^P \\ M^N & Z_m^{NP} \end{bmatrix}\begin{bmatrix} a_{bm}^{D*} \\ a_m^{P*} \end{bmatrix} + \begin{bmatrix} E_m^P \\ E_m^N \end{bmatrix} e_m + \begin{bmatrix} \epsilon^P \\ \epsilon^N + \phi_o^N + P_{bS}^N a_{bo}^{S*} \end{bmatrix}\right]$$

which after some rearrangement can be written as

$$\text{Var}\left[\begin{bmatrix} \mathbf{0} & \mathbf{0} & Z_o^P & \mathbf{0} \\ P_{bS}^N & P_{bD}^N & P^{NP} & I_{n_N} \end{bmatrix}\begin{bmatrix} a_{bo}^{S*} \\ a_{bo}^{D*} \\ a_o^{P*} \\ \phi_o^N \end{bmatrix} + \begin{bmatrix} M^P & Z_m^P \\ M^N & Z_m^{NP} \end{bmatrix}\begin{bmatrix} a_{bm}^{D*} \\ a_m^{P*} \end{bmatrix}\right] + E_m E_m' \sigma^2_{Em} + \text{Var}\begin{bmatrix} \epsilon^P \\ \epsilon^N \end{bmatrix} \qquad [B4]$$

Notice that the second row for the direct BV in [B4] is

$$\left[P_{bs}^N | P_{bD}^N\right]\begin{bmatrix} a_{bo}^{S*} \\ a_{bo}^{D*} \end{bmatrix} + P^{NP} a_o^{P*} + \phi_o^N = P_b^N a_{bo}^{N*} + P^{NP} a_o^{P*} + \phi_o^N = a_o^{N*}$$

The last equality follows from [A4] after replacing direct BV by starred BV (deviated from their means). Therefore, expressions [B3] and [B4] are equal.

APPENDIX C: Derivation of [14]: BV for Non-Parents

By [A4], BV of non-parents can be predicted by means of

$$\begin{bmatrix} \hat{a}_o^N \\ \hat{a}_m^N \end{bmatrix} = \left(I_2 \otimes \left[P_{bs}^N | P_{bD}^N\right]\right)\begin{bmatrix} \hat{a}_{bo}^N \\ \hat{a}_{bm}^N \end{bmatrix} + (I_2 \otimes P^{NP})\begin{bmatrix} \hat{a}_o^P \\ \hat{a}_m^P \end{bmatrix} + \begin{bmatrix} \hat{\phi}_o^N \\ \hat{\phi}_m^N \end{bmatrix} \qquad [C1]$$

After [11] is solved and the Mendelian residuals are calculated by means of [12] and [13], all predictors needed in [C1] are available except for the BV of the unknown sires of non-parents in a_{bo}^N and a_{bm}^N. To obtain the direct BV of these "phantom" sires write

$$a_{bo}^S = Q_{bS}^N g_o + a_{bo}^{s*} \qquad [C2]$$

where Q_{bo}^{NS} relates direct BV for missing sires (a_{bo}^s) of non-parents to direct genetic groups. The random variables with zero mean (a_{bo}^{S*}) in [C2] are the same as those in the residuals of non-parents in [10]. Hence, predictions for a_{bo}^{S*} are needed because estimates of g_o are available from [11]. These predictions can be obtained in the same way as the predictions for direct Mendelian residuals in [12]. Interestingly enough, for those non-parents with an unknown sire, both predictions are the same, i.e., BLUP(a_{bo}^{S*}) = BLUP(ϕ_o^N). To see that observe that BLUP for a_{bo}^{S*} and ϕ_o^N are linear functions of BLUP(ϵ^N) as both random variables are part of the residuals for non-parents in [10]. Therefore, if cov($a_{bo}^{S*}, \epsilon^{N'}$) = cov($\phi_o^N$, $\epsilon^{N'}$), their BLUP will be the same. To simplify notation the direct BV of the sire of non-parent i is denoted as a_{oSi}. Hence

$$\text{cov}(\boldsymbol{a}_{oSi}, \varepsilon_i^N) = \text{cov}(\boldsymbol{a}_{oSi}, \frac{1}{2}\boldsymbol{a}_{oSi}) = \frac{1}{2}\sigma^2_{Ao} = \text{cov}(\phi_{oi}^N, \epsilon_i^N) \quad [C3]$$

under the assumption of no-inbreeding. A similar proof can be used to show that BLUP for the starred maternal BV (with zero means) of the unknown sires of non-parents, BLUP($\boldsymbol{a}_{bm}^{S*}$), are equal to BLUP($\boldsymbol{\phi}_m^N$) and can be calculated by [13].

After taking estimators and predictors in [C2], replacing with them in [C1] and taking advantage of [C3], predictors for the BV of non-parents are

$$\begin{bmatrix} \hat{\boldsymbol{a}}_o^N \\ \hat{\boldsymbol{a}}_m^N \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}_{bo}^{NS}\hat{\boldsymbol{g}}_o + \hat{\phi}_o^N \\ \boldsymbol{Q}_{bm}^{NS}\hat{\boldsymbol{g}}_m + \hat{\phi}_m^N \end{bmatrix} + (\boldsymbol{I}_2 \otimes \boldsymbol{P}_{bD}^N) \begin{bmatrix} \hat{\boldsymbol{a}}_{bo}^{D*} \\ \hat{\boldsymbol{a}}_{bm}^{D*} \end{bmatrix} + (\boldsymbol{I}_2 \otimes \boldsymbol{I}^{NP}) \begin{bmatrix} \hat{\boldsymbol{a}}_o^P \\ \hat{\boldsymbol{a}}_m^P \end{bmatrix} + \begin{bmatrix} \hat{\phi}_o^N \\ \hat{\phi}_m^N \end{bmatrix}$$

or:

$$\begin{bmatrix} \hat{\boldsymbol{a}}_o^N \\ \hat{\boldsymbol{a}}_m^N \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}_{bo}^{NS} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{Q}_{bm}^{NS} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{g}}_o \\ \hat{\boldsymbol{g}}_m \end{bmatrix} + (\boldsymbol{I}_2 \otimes \boldsymbol{P}_{bD}^N) \begin{bmatrix} \hat{\boldsymbol{a}}_{bo}^{D*} \\ \hat{\boldsymbol{a}}_{bm}^{D*} \end{bmatrix} + (\boldsymbol{I}_2 \otimes \boldsymbol{P}^{NP}) \begin{bmatrix} \hat{\boldsymbol{a}}_o^P \\ \hat{\boldsymbol{a}}_m^P \end{bmatrix} + 2 \begin{bmatrix} \hat{\boldsymbol{\phi}}_o^N \\ \hat{\boldsymbol{\phi}}_m^N \end{bmatrix}$$

which is expression [14].