

I.N.R.A. C.R.J.

Département de Génétique Animale

BIBLIOTHÈQUE

F. 78352 JOUY EN JOSAS CEDEX

An expression of mixed animal model equations to account for different means and variances in the base population

Leopoldo Alfonso*, Joan Estany

Àrea de Producció Animal, Centre UdL-IRTA, 25198 Lleida, Spain

(Received 23 July 1998; accepted 28 January 1999)

Abstract – This paper presents a general expression to predict breeding values using animal models when the base population is selected, i.e. the means and variances of breeding values in the base generation differ among individuals. Rules for forming the mixed model equations are also presented. A numerical example illustrates the procedure. © Inra/Elsevier, Paris

mixed model equations / animal model / base population / selection

Résumé – **Expression générale des équations du modèle animal mixte tenant compte de la sélection des populations de base.** Cet article présente l'expression générale pour prédire les valeurs génétiques par le modèle animal quand la population de base est sélectionnée, c'est-à-dire quand cette population est un mélange de sous-populations à moyenne et variances génétiques différentes. On présente les règles de construction des équations du modèle mixte. La procédure est illustrée par un exemple numérique. © Inra/Elsevier, Paris

modèle mixte / modèle animal / population de base / sélection

1. INTRODUCTION

The prediction of breeding values involves assumptions on animals with unknown parents, commonly named the base animals. Correct understanding and definition of the base population are critical for animal models because all subsequent breeding values are tied to them. The usual assumption is to consider base animals unselected. However, this condition often does not hold

* Correspondence and reprints: Departamento de Producción Agraria, Universidad Pública de Navarra, Campus Arrosadía, 31006 Pamplona, Spain
E-mail: leo.alfonso@upna.es

because it is not always possible to trace the complete genealogy or to describe the selection process back to the unselected foundation generation. In this case, the distribution of breeding values is altered and, in particular, it is no longer valid to assume that the breeding values of the base animals have the same mean and variance, and that the genetic variance of the base generation is twice that of the Mendelian variance.

In a Gaussian setting, Henderson [6] derived a modification of the mixed model equations (MME) which led to the obtaining of predictors of breeding values that are unbiased even if base animals were selected, provided that the variance components associated with the model were known. In many applications these equations are difficult to set up and various alternatives have been suggested. In sire evaluation, Henderson [5] proposed to assign logically animals to fixed groups according to some existing prior knowledge of breeding values, or instead to treat animals as fixed if selection occurred in an unspecified manner. Quaas and Pollak [12] showed the equivalence between the MME for a sire model with genetic groups and those derived by Henderson [6] under his selection model, provided that the appropriate genetic groups were defined. The alternative formulation of the MME derived by Quaas and Pollak [12] for a model with genetic groups was exploited in Graser et al. [4] and in Quaas [11]. They gave easy rules to set up the equations corresponding to an animal model with base animals treated as fixed and an additive genetic animal model with groups and relationships, respectively. Cantet and Fernando [1] extended these rules to allow for heterogeneous additive genetic variances and segregation variance between groups. However, these rules assume that each base animal is randomly sampled within the group, and therefore that its variance is the same as before selection took place. Although Henderson [7, 8] and van der Werf and Thompson [14] developed MME that account for reduced genetic variance of base animals due to selection, they did not explicitly give a set of rules to set up the associated MME.

The purpose of this paper is to present a general approach to predicting breeding values when genetic means and variances of base animals are not homogeneous. The problem has been dealt with in the literature, and easy rules are available to set up the MME when individuals from the base population have different means [11] or can be easily derived when they have distinct genetic variances [14]. However, both aspects have not been dealt with in one practical approach. This paper brings these two problems together. The generalisation gives a convenient formulation for illustrating the relationships between several methods of dealing with selected base populations. This includes obtaining MME which can be constructed using an extension of the rules given by Quaas [11] to cope with different assumptions concerning the variance of breeding values of base animals. A numerical example is given.

2. THEORY

The usual animal model expression can be written as:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{a}_b + \mathbf{Z}_2\mathbf{a}_r + \mathbf{e}$$

where \mathbf{y} is the vector of records; \mathbf{b} is the vector of fixed effects; \mathbf{a}_b is the random vector of breeding values of base animals; \mathbf{a}_r is the random vector of breeding

values of non-base animals; \mathbf{e} is the random vector of residuals; and \mathbf{X} , \mathbf{Z}_1 and \mathbf{Z}_2 are known incidence matrices associated with \mathbf{b} , \mathbf{a}_b and \mathbf{a}_r , respectively.

The vector of breeding values of non-base animals can be partitioned as:

$$\mathbf{a}_r = \mathbf{Q}\mathbf{a}_b + \mathbf{s}^*$$

where \mathbf{s}^* is a linear transformation of the random vector of the Mendelian sampling effects (\mathbf{s}) of animals with known parents, such that

$$\mathbf{s}^* = (\mathbf{I} - \mathbf{P}_2)^{-1}\mathbf{s} \quad (1)$$

where \mathbf{P}_2 is a matrix relating non-base animals among themselves; and \mathbf{Q} is the incidence matrix relating base animals with their descendants, such that

$$\mathbf{Q} = (\mathbf{I} - \mathbf{P}_2)^{-1}\mathbf{P}_1 \quad (2)$$

where \mathbf{P}_1 is a matrix relating base animals with non-base animals. \mathbf{P}_1 and \mathbf{P}_2 are matrices with 0.5 in the parent's columns in each row.

We can then write [4]:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + [\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{Q}]\mathbf{a}_b + \mathbf{Z}_2\mathbf{s}^* + \mathbf{e} \quad (3)$$

With this model (3), the expectation and dispersion matrices for the vector of observations are:

$$\mathbf{E}(\mathbf{y}) = \mathbf{X}\mathbf{b} + (\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{Q})\mathbf{E}(\mathbf{a}_b)$$

and

$$\begin{aligned} \mathbf{V}(\mathbf{y}) = & [\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{Q}]\mathbf{V}(\mathbf{a}_b)[\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{Q}]' + \mathbf{Z}_2\mathbf{V}(\mathbf{s}^*)\mathbf{Z}_2' + \\ & [\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{Q}]\text{Cov}(\mathbf{a}_b, \mathbf{s}^*)\mathbf{Z}_2' + \mathbf{Z}_2\text{Cov}(\mathbf{a}_b, \mathbf{s}^*)[\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{Q}]' + \mathbf{V}(\mathbf{e}) \end{aligned}$$

Note that as Mendelian sampling effects are independent of ancestral breeding values we have

$$\text{Cov}(\mathbf{a}_b, \mathbf{s}^*) = \text{Cov}(\mathbf{a}_b, (\mathbf{I} - \mathbf{P}_2)^{-1}\mathbf{s}) = (\mathbf{I} - \mathbf{P}_2)^{-1}\text{Cov}(\mathbf{a}_b, \mathbf{s}) = \mathbf{0}$$

Further, it can be shown that the dispersion matrix of Mendelian sampling, $\text{var}(\mathbf{s}) = \mathbf{H}_0$, is diagonal with the i th diagonal element defined as:

$$\text{var}(a_i) - \frac{1}{4} [\text{var}(a_j) + \text{var}(a_k)] - \frac{1}{2} \text{cov}(a_j, a_k)$$

where j and k are the parents of i .

Thus, following equation (1), we can write the variance of \mathbf{s}^* as:

$$\mathbf{V}(\mathbf{s}^*) = \mathbf{S} = (\mathbf{I} - \mathbf{P}_2)^{-1}\mathbf{H}_0 [(\mathbf{I} - \mathbf{P}_2)^{-1}]' \quad (4)$$

We will denote that $\mathbf{V}(\mathbf{e}) = \mathbf{R}$. To complete the definition of the model, we need only to specify the expectation and dispersion matrices for \mathbf{a}_b . This will serve to develop different hypotheses about the mean and variance of the base population. In most mixed models either the mean or the variance is assumed to be zero. Hence, to be general, we will consider that $\mathbf{E}(\mathbf{a}_b) = \mathbf{Q}_b\mathbf{g}$ and $\mathbf{V}(\mathbf{a}_b) = \mathbf{H}_b$, where \mathbf{g} is the vector of base population means, \mathbf{Q}_b an incidence matrix relating the base animals to their respective groups, and \mathbf{H}_b is the dispersion matrix of breeding values of base animals.

Expression (3) can be rewritten as:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + [\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{Q}]\mathbf{Q}_b\mathbf{g} + [\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{Q}]\mathbf{a}_b^* + \mathbf{Z}_2\mathbf{s}^* + \mathbf{e}$$

With this model the vector of breeding values of base animals is:

$$\mathbf{a}_b = \mathbf{Q}_b\mathbf{g} + \mathbf{a}_b^*$$

Now, following the modification of Quaas and Pollak [12] in a similar manner to that described by Graser et al. [4], the associated MME are:

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{0} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_1 & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}_2 \\ \mathbf{0} & \mathbf{Q}'_b\mathbf{H}_b^{-1}\mathbf{Q}_b & -\mathbf{Q}'_b\mathbf{H}_b^{-1} & \mathbf{0} \\ \mathbf{Z}'_1\mathbf{R}^{-1}\mathbf{X} & -\mathbf{H}_b^{-1}\mathbf{Q}_b & \mathbf{Z}'_1\mathbf{R}^{-1}\mathbf{Z}_1 + \mathbf{H}_b^{-1} + \mathbf{Q}'\mathbf{S}^{-1}\mathbf{Q} & -\mathbf{Q}'\mathbf{S}^{-1} \\ \mathbf{Z}'_2\mathbf{R}^{-1}\mathbf{X} & \mathbf{0} & -\mathbf{S}^{-1}\mathbf{Q} & \mathbf{Z}'_2\mathbf{R}^{-1}\mathbf{Z}_2 + \mathbf{S}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{g}} \\ \hat{\mathbf{a}}_b \\ \hat{\mathbf{a}}_r \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{0} \\ \mathbf{Z}'_1\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'_2\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \quad (5)$$

Absorption of the equations for the genetic groups ($\hat{\mathbf{g}}$), and using equations (2) and (4), permit us to rewrite the MME in equation (5) as

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} \\ \begin{bmatrix} \mathbf{Z}'_1 \\ \mathbf{Z}'_2 \end{bmatrix} \otimes \mathbf{R}^{-1}\mathbf{X} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \begin{bmatrix} \mathbf{Z}'_1 \\ \mathbf{Z}'_2 \end{bmatrix} \otimes \mathbf{R}^{-1}\mathbf{y} \end{bmatrix} + \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1} \otimes [\mathbf{Z}_1\mathbf{Z}_2] \\ \begin{bmatrix} \mathbf{Z}'_1\mathbf{R}^{-1}\mathbf{Z}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}'_2\mathbf{R}^{-1}\mathbf{Z}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{P}'_1\mathbf{H}_0^{-1}\mathbf{P}_1 & \mathbf{P}'_1\mathbf{H}_0^{-1}(\mathbf{P}_2 - \mathbf{I}) \\ (\mathbf{P}_2 - \mathbf{I})'\mathbf{H}_0^{-1}\mathbf{P}_1 & (\mathbf{I} - \mathbf{P}_2)'\mathbf{H}_0^{-1}(\mathbf{I} - \mathbf{P}_2) \end{bmatrix} + \begin{bmatrix} \mathbf{H}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \end{bmatrix}$$

where

$$\mathbf{H}^{-1} = \mathbf{H}_b^{-1} \left(\mathbf{I} - \mathbf{Q}_b (\mathbf{Q}'_b\mathbf{H}_b^{-1}\mathbf{Q}_b)^{-1} \mathbf{Q}'_b \right) \mathbf{H}_b^{-1} \quad (6)$$

and $\hat{\mathbf{a}}$ is the vector of breeding values of base ($\hat{\mathbf{a}}_b$) and non-base animals ($\hat{\mathbf{a}}_r$).

Now, calling

$$\mathbf{G}^{*-1} = \text{Var} \begin{bmatrix} \mathbf{a}_b \\ \mathbf{s}^* \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{P}'_1\mathbf{H}_0^{-1}\mathbf{P}_1 & \mathbf{P}'_1\mathbf{H}_0^{-1}(\mathbf{P}_2 - \mathbf{I}) \\ (\mathbf{P}_2 - \mathbf{I})'\mathbf{H}_0^{-1}\mathbf{P}_1 & (\mathbf{I} - \mathbf{P}_2)'\mathbf{H}_0^{-1}(\mathbf{I} - \mathbf{P}_2) \end{bmatrix} + \begin{bmatrix} \mathbf{H}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

and $\mathbf{Z} = [\mathbf{Z}_1 \quad \mathbf{Z}_2]$ the prediction of breeding values when base animals are selected is then obtained by solving the following MME:

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{*-1} \end{bmatrix} \cdot \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{a}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \quad (7)$$

The calculation of \mathbf{G}^{*-1} is simplified if all the groups are assumed to have the same additive genetic variance, and base animals are unrelated and non-inbred, because in that case \mathbf{G}^{*-1} is the usual inverse of the relationship matrix. Otherwise, the calculation of \mathbf{G}^{*-1} requires computing \mathbf{H}_0 introducing the segregation variance between groups and inbreeding, though these effects can be easily accommodated using for example the algorithms given by Cantet et al. [1] and Meuwissen and Luo [10], respectively.

The second term of \mathbf{G}^{*-1} requires the computation of \mathbf{H}^{-1} . However, from inspection of MME in equation (5) it can be seen that, if no inbreeding is assumed and base animals are genetically unrelated, \mathbf{H}^{-1} does not need to be calculated because \mathbf{G}^{*-1} can be constructed directly by extending the algorithm of Quaas [11]. In particular, if base animals are sampled at random from some selected populations, and, for simplicity, are assumed to be genetically unrelated, then \mathbf{H}_b is diagonal with the i th diagonal element defined as $\delta_i \sigma_a^2$, where δ_i accounts for the reduction in the genetic variance σ_a^2 due to selection. In this case, $\mathbf{G}^{*-1} = \mathbf{A}^{*-1}(1/\sigma_a^2)$ and \mathbf{A}^{*-1} can be computed, for $m =$ number of unknown parents of an individual, replacing $x (= 4/[k+2])$ in the rules of Quaas [11] with:

- $x = 2$, if $m = 0 (k = m)$;
- $x = [4/(2 + \delta_j)]$, if $m = 1$ and the unknown parent is from a population with variance $\delta_j \sigma_a^2 (k = \delta_j)$;
- and $x = [4/(2 + \delta_j + \delta_k)]$, if $m = 2$ and the unknown parents are from populations with variance $\delta_j \sigma_a^2$ and $\delta_k \sigma_a^2 (k = \delta_j + \delta_k)$.

3. NUMERICAL EXAMPLE

Consider the following pedigree:

Animal	Sire	Dam
1	-	-
2	-	-
3	-	-
4	-	-
5	1	2
6	1	4
7	3	2
8	3	4
9	-	4
10	3	4
11	6	2

All base sires and dams come from the same population. Dams were taken at random and sires were selected from the offspring of the 1% of the phenotypically best animals.

Records were made in two time periods as follows:

Animal	Time	Record
5	2	8
6	1	10
7	1	12
8	1	9
9	2	7
10	1	11
11	2	5

Variance components in original populations were: $\sigma_p^2 = 16$, $\sigma_a^2 = 4$.

Two different genetic groups can be defined: g_1 for the selected base males and g_2 for the randomly chosen base females, both with different additive genetic means and variances. Assigning a hypothetical base animal to these groups (g_1 and g_2) genetic groups can be treated as fixed effects ($\delta = \infty$).

Selection carried out in males is known. Therefore, assuming normality, the proportion of genetic additive variance after selection (δ) can be derived from the following expression: $\delta = 1 - i(i-w)h^2$, with i being the selection intensity value, w the standardised truncation point value and h^2 the heritability value [13].

Following the proposed rules, we have:

Animal	δ_{sire}	δ_{dam}	x_i
g_1	∞	∞	0
g_2	∞	∞	0
1	0.776	0.776	1.126
2	1	1	1
3	0.776	0.776	1.126
4	1	1	1
5	1	1	2
6	1	1	2
7	1	1	2
8	1	1	2
9	0.776	1	1.441
10	1	1	2
11	1	1	2

The associated MME are:

$$\begin{bmatrix}
 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 \\
 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\
 0 & 0 & 7.837 & 0 & -3.378 & 0 & -3.378 & 1.081 & 0 & 0 & 0 & 0 & -2.161 & 0 & 0 \\
 0 & 0 & 0 & 6 & 0 & -3 & 0 & -3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & -3.378 & 0 & 6.378 & 1.5 & 0 & 1.5 & -3 & -3 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & -3 & 1.5 & 7.5 & 1.5 & 0 & -3 & 1.5 & -3 & 0 & 0 & 0 & -3 \\
 0 & 0 & -3.378 & 0 & 0 & 1.5 & 7.878 & 3 & 0 & 0 & -3 & -3 & 0 & -3 & 0 \\
 0 & 0 & 1.081 & -3 & 1.5 & 0 & 3 & 8.581 & 0 & -3 & 0 & -3 & -2.161 & -3 & 0 \\
 0 & 1 & 0 & 0 & -3 & -3 & 0 & 0 & 7 & 0 & 0 & 0 & 0 & 0 & 0 \\
 1 & 0 & 0 & 0 & -3 & 1.5 & 0 & -3 & 0 & 8.5 & 0 & 0 & 0 & 0 & -3 \\
 1 & 0 & 0 & 0 & 0 & -3 & -3 & 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 \\
 1 & 0 & 0 & 0 & 0 & 0 & -3 & -3 & 0 & 0 & 0 & 7 & 0 & 0 & 0 \\
 0 & 1 & -2.161 & 0 & 0 & 0 & 0 & -2.161 & 0 & 0 & 0 & 0 & 5.323 & 0 & 0 \\
 1 & 0 & 0 & 0 & 0 & 0 & -3 & -3 & 0 & 0 & 0 & 0 & 0 & 7 & 0 \\
 0 & 1 & 0 & 0 & 0 & 0 & -3 & 0 & 0 & -3 & 0 & 0 & 0 & 0 & 7
 \end{bmatrix}
 \begin{bmatrix}
 \hat{b}_1 \\
 \hat{b}_2 \\
 \hat{g}_1 \\
 \hat{g}_2 \\
 \hat{a}_1 \\
 \hat{a}_2 \\
 \hat{a}_3 \\
 \hat{a}_4 \\
 \hat{a}_5 \\
 \hat{a}_6 \\
 \hat{a}_7 \\
 \hat{a}_8 \\
 \hat{a}_9 \\
 \hat{a}_{10} \\
 \hat{a}_{11}
 \end{bmatrix}
 =
 \begin{bmatrix}
 42 \\
 20 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 0 \\
 8 \\
 10 \\
 12 \\
 9 \\
 7 \\
 11 \\
 5
 \end{bmatrix}$$

The coefficient matrix has order 15 but rank 14. Imposing the restriction $\hat{g}_1 = 0$ the solution is: $\hat{b} = (15.622, 12.593)$, $\hat{g} = (0, -10.162)$ and $\hat{a} = (-0.006, -9.921, 0.053, -10.402, -4.911, -5.264, -4.747, -5.381, -5.275, -5.096, -7.593)$. The large difference estimated between groups can be explained by the higher proportionate contribution of the group of females to the records made in the second time period.

For the usual genetic group model, assuming that additive genetic variance is the same for both groups of base animals, males and females, there is also one dependency in the equations. Imposing the same restriction ($\hat{g}_1 = 0$) the solution is: $\hat{b} = (15.626, 12.602)$, $\hat{g} = (0, -10.174)$ and $\hat{a} = (-0.007, -9.934, 0.059, -10.414, -4.918, -5.270, -4.750, -5.384, -5.286, -5.099, -7.602)$.

MME for the simplest animal model, assuming $E(a_b) = 0$ and $V(a_b) = I \sigma_a^2$, have full rank. The solution is: $\hat{b} = (10.586, 6.701)$ and $\hat{a} = (0.009, 0.133, 0.050, -0.249, 0.246, -0.286, 0.280, -0.312, -0.040, -0.026, -0.309)$.

4. DISCUSSION

The results presented in this paper permit us to obtain a general expression to predict breeding values using animal models when the means and variances of breeding values in the base generation differ among individuals. This can be accomplished using equation (5) or equation (7) with a proper definition of \mathbf{H} in equation (6). In particular, it is through \mathbf{Q}_b and \mathbf{H}_b that we account for the distribution of breeding values of base animals, and can illustrate the correspondence among different models for selected base populations. Thus, if $\mathbf{Q}_b = \mathbf{0}$ and $\mathbf{H}_b = \mathbf{I} \sigma_a^2$, the expression (7) leads us to the habitual MME under a non-selection model. With $\mathbf{H}_b^{-1} = \mathbf{0}$, which can be obtained by setting $\delta_i = \infty$, we can represent the genetic groups model [11] or the fixed base animal model [4], depending on whether $\mathbf{Q}_b \neq \mathbf{0}$ or $\mathbf{Q}_b = \mathbf{0}$, respectively.

Similarly, the MME described in van der Werf and Thompson [14] are the same as in equation (7) with $\mathbf{H}_b^{-1} = \delta \mathbf{I} (1/\sigma_a^2)$ and $\mathbf{Q}_b = \mathbf{0}$. Further, when selection can be described as a linear function of breeding values of base animals ($\mathbf{M}'\mathbf{a}_b$), it can be shown that equation (7) is equivalent to equation (3) in Henderson [8] when $\mathbf{M} = \mathbf{H}_b^{-1} \mathbf{Q}_b$, and, therefore, $\mathbf{Q}_b' \mathbf{H}_b^{-1} \mathbf{a}_b$ represents the conditional variable upon which selection is assumed to be based. This can be interpreted generally as a weighted grouping, where groups are weighted by the dispersion matrix of breeding values of base animals. Alternatively, the results of Famula [2] serve to show that this is equivalent to a model of restricted selection using $\mathbf{H}_b^{-1} \mathbf{Q}_b$ as a restriction matrix.

Hence, predictions of \mathbf{a}_b deviations from their group mean are independent of selection decisions made in the past and, assuming normality, selection can be ignored. Note, however, that this is not true if descendants of base animals are also selected, unless they are selected on linear, translation invariant functions of the observations [6]. The latter condition would not be satisfied when the selection criterion included the group effect or, more generally, when base animals were treated as fixed [14]. Nonetheless, this condition for ignoring selection does not need to be met when likelihood [9] or Bayesian [3] methods of inference are used, and it has not been demonstrated that this property leads to maximising the expected genetic progress, as Fernando and Gianola [3] have shown in a simulated example.

Equation (7) can also be useful in the estimation of variance components when, as in the example presented, selection can be simply modelled. In this case, the problem of selected base animals could be reduced to estimating some extra parameters, although the amount and the structure of available data would condition the reliability of estimates [14].

5. CONCLUSION

When additive genetic means and variances of base animals are not homogeneous, prediction of breeding values can be obtained by means of animal models if the covariance matrix of additive genetic values is properly defined. MME construction is similar to that with homogeneous mean and variance in the base population. The different methods that have been proposed for prediction of breeding values when base population animals have been selected in some non-random manner can be deduced from a general expression of MME.

ACKNOWLEDGEMENT

This research was supported by the CICYT Research Project AGF94-1016 of the Ministerio de Educación y Ciencia, Spain.

REFERENCES

- [1] Cantet R.J.C., Fernando R.L., Prediction of breeding values with additive animal models for crosses from two populations, *Genet. Sel. Evol.* 27 (1995) 323-334.

- [2] Famula T.R., An equivalence between models of restricted selection and genetic groups, *Theor. Appl. Genet.* 71 (1985) 413–416.
- [3] Fernando R.L., Gianola D., Statistical inferences in populations undergoing selection or non-random mating, in: Gianola D., Hammond K. (Eds.), *Advances in Statistical Methods for Genetic Improvement of Livestock*, Springer-Verlag, Berlin, 1990, pp. 437–453.
- [4] Graser H.-U., Smith S.P., Tier B., A derivative-free approach for estimating variance components in animal models by restricted maximum likelihood, *J. Anim. Sci.* 64 (1987) 1362–137.
- [5] Henderson C.R., Sire evaluation and genetic trends, in: *Proceedings of the Animal Breeding and Genetics Symposium in Honor of Dr J.L. Lush, ASAS and ADSA*, Champaign, 1973, pp. 10–41.
- [6] Henderson C.R., Best linear unbiased estimation and prediction under a selection model, *Biometrics* 31 (1975) 423–447.
- [7] Henderson C.R., Best linear unbiased prediction using relationship matrices derived from selected base populations, *J. Dairy Sci.* 68 (1985) 443–448.
- [8] Henderson C.R., A simple method to account for selected base population, *J. Dairy Sci.* 71 (1988) 3399–3404.
- [9] Im S., Fernando R.L., Gianola D., Likelihood inferences in animal breeding under selection: a missing-data theory viewpoint, *Genet. Sel. Evol.* 21 (1989) 399–414.
- [10] Meuwissen T.H.E., Luo Z., Computing inbreeding coefficients in large populations, *Genet. Sel. Evol.* 24 (1992) 305–313.
- [11] Quaas R.L., Additive genetic model with groups and relationships, *J. Dairy Sci.* 71 (1988) 1338–1345.
- [12] Quaas R.L., Pollack E.J., Modified equations for sire models with groups, *J. Dairy Sci.* 64 (1981) 1868–1872.
- [13] Robertson A., The effect of selection on the estimation of genetic parameters, *Z. Tierzüchtg. Züchtgsbiol* 94 (1977) 131–135.
- [14] Van der Werf J., Thompson R., Variance decomposition in the estimation of genetic variance with selected data, *J. Anim. Sci.* 70 (1992) 2975–2985.