ANALYSIS OF SELECTION EXPERIMENTS USING MIXED MODEL METHODOLOGY

D. A. Sorensen and B. W. Kennedy
National Institute of Animal Science
Orum Sdr., Denmark
and
University of Guelph
Ontario, Canada N1G 2W1

ABSTRACT

Use of mixed model techniques to estimate genetic variance and selection response is illustrated by simple examples. A minimum variance quadratic unbiased estimator (MIVQUE) of genetic variance using a reduced animal model is derived. Properties of the mixed model estimator of response are discussed and illustrated with results from Monte Carlo simulation. The mixed model estimator of response requires knowledge of the base population heritability. When the latter is not known, simulation results suggest that using a MIVQUE estimate obtained from the data yields estimates of response in good agreement with the true response. If a number of conditions are satisfied, the mixed model estimator of response partitions the phenotypic trend into its genetic and environmental components, without need for a control population. These conditions are unlikely to hold in long-term selection experiments. More work is needed to understand the implications of finite numbers of loci or the presence of unaccounted natural selection opposing artificial selection, for example, on the properties of the mixed model estimator of response. (Key Words: Selection Responses, Mixed Model Methods, Best Linear Unbiased Predictor.)

Introduction

A first step in the interpretation of the results of selection experiments is the estimation of genetic parameters such as direct and correlated responses to selection, heritabilities and genetic correlations. In a simply designed experiment with non-overlapping generations, response to selection may be estimated using least-squares procedures as the phenotypic mean of the offspring of selected parents. Environmental effects are accounted for by the use of an unselected control. The data from the experiment also can be used to obtain estimates of heritabilities using simple regression-type estimators such as cumulative response divided by cumulative selection differential.

Alternatively, a mixed model procedure can be used to analyse the data to yield estimates of genetic variance and of selection response in a single analysis. Mixed model type estimators are demanding computationally, but they have a number of well-defined statistical properties that the experimenter may find desirable. In this paper, we illustrate with simple examples the use of mixed model techniques in the analysis of data from selection experiments. Some properties of mixed model type estimators are discussed and illustrated using Monte Carlo simulation. The emphasis of the paper is on analysis of data; the important aspect of design is not included.

Estimation of Genetic Variance

In this section we illustrate the use of two estimators of genetic variances: one is a modified Method 3 estimator (Henderson, 1953) and the other is a minimum variance quadratic unbiased estimator (MIVQUE). We also develop and illustrate the use of MIVQUE on a model that Qaaas and Pollak (1980) termed the reduced animal model or gametic model.
Assume that data from a control population spanning several generations are available. Let the mixed model describing the data be as follows:

\[ y = Xb + Zu + e \]  

where \( y \) denotes the \( n \times 1 \) vector of observations, \( b \) is a vector of fixed effects, \( u \) is a random vector of sire effects, \( e \) is a vector of random errors and \( X \) and \( Z \) are design matrices. It is usually assumed that

\[ \text{Var} (u) = \sigma_u^2, \quad \text{Var} (e) = \sigma_e^2 \text{ and } \text{Var} (y) = ZZ'\sigma_u^2 + I\sigma_e^2, \]  

(2)

where \( \sigma_u^2 \) is the variance component between half-sib progeny groups, \( \sigma_e^2 \) is the error variance and \( I \) is the identity matrix. Expression (2) assumes that sires are unrelated and that the only covariances are those within half-sib groups. However, as random drift develops with each cycle of random matings, this assumption does not hold. To acknowledge the additional correlated structure among the observations, the correct representation is

\[ \text{Var} (y) = ZAZ'\sigma_a^2 + I\sigma_e^2 \]  

(3)

where \( A \) is the \( n \times n \) matrix of additive genetic relationships, \( \sigma_a^2 \) is the base population additive genetic variance and \( \sigma_e^2 \) is the environmental variance (assumed constant each generation). To obtain unbiased estimates of \( \sigma_a^2 \) and of \( \sigma_e^2 \), Henderson’s (1953) method 3 estimator was modified as follows (Sorensen and Kennedy, 1982):

\[
\begin{bmatrix}
\sigma_a^2 \\
\sigma_e^2
\end{bmatrix} = \frac{1}{\sum_{i=1}^{n} F_i - \text{tr} (Z(Z'Z)^{-1}Z'A)} 
\begin{bmatrix}
\text{tr}(Z(Z'Z)^{-1}Z'A) - \text{tr}(X(X'X)^{-1}X'A) \\
\text{tr}(Z(Z'Z)^{-1}Z'A)
\end{bmatrix}^{-1} 
\begin{bmatrix}
R(u/b) \\
\text{SSE}
\end{bmatrix},
\]  

(4)

where

- \( t \) is the number of generations,
- \( p \) is the number of sires within generations,
- \( n \) is the total number of records,
- \( F_i \) is the inbreeding coefficient of the \( i \)th individual,
- \( R(u/b) = R(u,b) - R(b) \) is the reduction in sums of squares due to fitting sires after the fixed effects and
- \( \text{SSE} = y'y - R(u,b) \) is the error sum of squares.

Essentially with this estimator, the expectations of \( R(u/b) \) and of \( \text{SSE} \) are taken using the correct variance structure of the data, specified by (3).

We illustrate the method with a simple example. Consider the following data:

<table>
<thead>
<tr>
<th>Animal</th>
<th>Sire</th>
<th>Dam</th>
<th>Record</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Unknown</td>
<td>Unknown</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>

Animals 1, 2, 3 and 4 were sampled at random from a base population. Animals 1, 2 and 3 produced progeny. The matrix of relationships among the progeny for this example is:

\[
A = \begin{bmatrix} 1 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1 & 0 & 0 \\ 1/4 & 0 & 1 & 1/2 \\ 1/4 & 0 & 1/2 & 1 \end{bmatrix}
\]

The usual model to estimate the additive genetic variance from the progeny data using a sire model is

\[
y_{ij} = \mu + s_i + e_{ij},
\]

where \( y_{ij} \) is the record, \( \mu \) is the mean, \( s_i \) is the effect of the \( i \)th sire and \( e_{ij} \) is a random error. In matrix form this is

\[
y = \mathbf{1} \mu + Zs + e
\]

where \( \mathbf{1} \) is a column vector of ones, \( Z \) is an incidence matrix, \( s \) is the vector of sire effects and \( e \) is the vector of error terms. It is usually assumed that
We compute the following reductions:

\[ \begin{align*}
R(\mu/s) &= R(\mu,s) - R(\mu) = 53 - 49 = 4 \\
SSE &= y'y - R(\mu,s) = 54 - 53 = 1
\end{align*} \]

The expectations of these reductions are taken under the correct variance structure of \( y \), which for the example is

\[ \text{Var}(y) = \lambda_0^2 + \lambda_1 e^2 \]

where \( \lambda_0^2 \) and \( \lambda_1 e^2 \) are the additive genetic variance in the base population and the environmental variance, respectively. Notice that the relationship matrix takes care of all the covariances among the progeny observations, while in the sire model, \( \lambda_0^2 \) only allows for relationships within half-sib groups and assumes (incorrectly) that half-sib groups in this example are uncorrelated. Taking expectations we arrive at the system of equations (4) which for our example are:

\[ \begin{bmatrix}
\lambda_0^2 \\
\lambda_1^2
\end{bmatrix} = \begin{bmatrix}
2.75 - 1.625 \\
4 - 2.75
\end{bmatrix}^{-1} \begin{bmatrix}
4 \\
73.875 \\
(4 - 2.75) \\
4 - (1)(2)
\end{bmatrix} \]

This modified method 3 estimator is unbiased and translation invariant. However, the estimator does not use all the data; only the part of the relationship matrix comprising block diagonals of relationships among observations within a generation is used. Relationships among non-collateral relatives are ignored. The omission of relationships across generations has important implications in the analyses of data that have been generated by selection.

An alternative estimator that makes use of all the data is a generalization of Rao's (1971) minimum variance quadratic unbiased estimator (MIVQUE), which can be found in Henderson (1984). This estimator has the desirable properties associated with maximum likelihood type estimators, but it has the disadvantage that prior variance ratios must be provided. With random mating (no selection), if good-enough prior estimates are available, sampling variances are smaller than for any other quadratic unbiased estimator. Use of an incorrect prior still leads to unbiased estimates of variance components, though not minimum variance, for unselected populations. The need for knowledge of the correct prior can be relaxed at additional computing cost, using iterative procedures such as iterative MIVQUE or restricted maximum likelihood (REML) of the type developed by Thompson (1977). In fact, ignoring complications caused by constraints on the parameter space or by nonconvergence or convergence, iterative MIVQUE is identical to REML (Harville, 1977).

To illustrate the use of MIVQUE, we review results of Sorensen and Kennedy (1984b). Consider the mixed model:

\[ y = Xb + Za + \xi \]  

The elements of (5) are the same as in (1) except for the random vectors \( a \) and \( \xi \) which represent additive genetic values and environmental effects of individual animals respectively. With one record per individual, \( Z \) in (5) is equal to \( I \), the identity matrix. As in (3) we assume

\[ \text{Var}(y) = ZAZ'\lambda_0^2 + \lambda_1 e^2 \]

where \( \lambda_0^2 \) and \( \lambda_1 e^2 \) are the additive genetic variance in the base population and the environmental variance, respectively. The vectors \( b \) and \( a \) are estimated from (7):

\[ \begin{bmatrix}
b \\
a
\end{bmatrix} = \begin{bmatrix}
X'X & X'Z \\
Z'X & Z'Z + A^{-1}\gamma
\end{bmatrix}^{-1} \begin{bmatrix}
X'y \\
Z'y
\end{bmatrix} \]

where \( \gamma \) is the a priori value of \( \lambda_1^2 / \lambda_0^2 \). MIVQUE of \( \lambda_0^2 \) and of \( \lambda_1 e^2 \) are obtained as follows:

\[ \begin{bmatrix}
\lambda_0^2 \\
\lambda_1^2
\end{bmatrix} = \begin{bmatrix}
q - 2\gamma tr(A^{-1}C_{11}) + \gamma^2 tr(A^{-1}C_{11})^2 \\
q - \gamma^2 tr(A^{-1}C_{11})
\end{bmatrix}^{-1} \begin{bmatrix}
Q_1 \\
Q_2
\end{bmatrix} \]

where

\[ \begin{bmatrix}
Q_1 \\
Q_2
\end{bmatrix} = \begin{bmatrix}
tr(A^{-1}C_{11}) - \gamma tr(A^{-1}C_{11})^2 \\
tr(A^{-1}C_{11}) - q + tr(A^{-1}C_{11})^2
\end{bmatrix} \]

and

\[ \begin{bmatrix}
\lambda_0^2 \\
\lambda_1^2
\end{bmatrix} = \begin{bmatrix}
q - 2\gamma tr(A^{-1}C_{11}) + \gamma^2 tr(A^{-1}C_{11})^2 \\
q - \gamma^2 tr(A^{-1}C_{11})
\end{bmatrix}^{-1} \begin{bmatrix}
Q_1 \\
Q_2
\end{bmatrix} \]

and

\[ \begin{bmatrix}
\lambda_0^2 \\
\lambda_1^2
\end{bmatrix} = \begin{bmatrix}
q - 2\gamma tr(A^{-1}C_{11}) + \gamma^2 tr(A^{-1}C_{11})^2 \\
q - \gamma^2 tr(A^{-1}C_{11})
\end{bmatrix}^{-1} \begin{bmatrix}
Q_1 \\
Q_2
\end{bmatrix} \]

and

\[ \begin{bmatrix}
\lambda_0^2 \\
\lambda_1^2
\end{bmatrix} = \begin{bmatrix}
q - 2\gamma tr(A^{-1}C_{11}) + \gamma^2 tr(A^{-1}C_{11})^2 \\
q - \gamma^2 tr(A^{-1}C_{11})
\end{bmatrix}^{-1} \begin{bmatrix}
Q_1 \\
Q_2
\end{bmatrix} \]
where \( r(X) \) is the rank of \( X \), \( q \) is the number of elements in \( a \), \( n \) is the total number of records, \( Q_1 = \hat{a}'A^{-1}\hat{a} \) and \( Q_2 = y'y - \hat{b}'X'y - \hat{a}'Z'y \).

Using Monte Carlo simulation, Sorensen and Kennedy (1984b) have shown the MIVQUE estimator has considerably smaller sampling variance than the modified method 3 estimator. We illustrate MIVQUE with the data used previously to illustrate method 3. MIVQUE, however, makes use of all the data, parental as well as progeny. The matrix of relationships among all eight animals and its inverse for this example are:

\[
A = \begin{bmatrix}
1 & 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 \\
1 & 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 \\
Symmetric & 1 & 0 & 1/2 & 0 & 1/4 & 1/4 & 1/4 \\
& 1 & 1/4 & 1/4 & 1/4 & 1/2 & 0 & 0 \\
& 1 & 0 & 0 & 0 & 1 & 1/2 & 1 \\
\end{bmatrix}
\]

\[
A^{-1} = \begin{bmatrix}
2 & 0 & 1/2 & 1/2 & -1 & -1 & 0 & 0 \\
2 & 1 & 0 & 0 & 0 & -1 & -1 & 0 \\
2 & 1/2 & 0 & -1 & 0 & -1 & -1 & 0 \\
Symmetric & 2 & 0 & 0 & 0 & 2 & 0 & 0 \\
& 2 & 0 & 0 & 2 & 0 & 0 & 2 \\
\end{bmatrix}
\]

The model is given by (5) and (6) with \( Z = 1 \) because there is one record per individual. We assume the prior value of \( \sigma_e^2/\sigma_a^2 \) is 1 (i.e., \( h^2 = .5 \)). The mixed model equations (7) are:

\[
\begin{bmatrix}
8 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 8 \\
3 & 0 & 1/2 & 1/2 & -1 & -1 & 0 & 0 & 3 \\
3 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 3 \\
3 & 1/2 & 0 & 0 & -1 & 0 & -1 & -1 & 3 \\
2 & 1/2 & 0 & -1 & 0 & -1 & 0 & 0 & 2 \\
Symmetric & 3 & 0 & 0 & 0 & 3 & 0 & 0 & 3 \\
& 3 & 0 & 0 & 3 & 0 & 0 & 3 & 3 \\
\end{bmatrix}
\begin{bmatrix}
\hat{\mu} \\
\hat{\alpha}_1 \\
\hat{\alpha}_2 \\
\hat{\alpha}_3 \\
\hat{\alpha}_4 \\
\hat{\alpha}_5 \\
\hat{\alpha}_6 \\
\hat{\alpha}_7 \\
\hat{\alpha}_8 \\
\end{bmatrix}
= \begin{bmatrix}
28 \\
2 \\
2 \\
3 \\
7 \\
3 \\
2 \\
4 \\
5 \\
\end{bmatrix}
\]
The unique solution is (approximately):

\[
\begin{bmatrix}
\beta \\
\hat{\alpha}_1 \\
\hat{\alpha}_2 \\
\hat{\alpha}_3 \\
\hat{\alpha}_4 \\
\hat{\alpha}_5 \\
\hat{\alpha}_6 \\
\hat{\alpha}_7 \\
\end{bmatrix} =
\begin{bmatrix}
.40 & -.25 & -.27 & -.23 & -.31 & -.29 & -.31 & 28 \\
.59 & .16 & .14 & .11 & .33 & .32 & .18 & 2 \\
.59 & .11 & .14 & .17 & .18 & .32 & .32 & 2 \\
.59 & .16 & .34 & .19 & .32 & .32 & .32 & 3 \\
.59 & .16 & .31 & .18 & .18 & .66 & .27 & .27 & 7 \\
& .64 & .22 & .22 & .65 & .32 & .32 & 4 \\
& & & & .363 & .65 & .65 & 5 \\
\end{bmatrix}
\]

It can be easily verified that \(\text{tr}(A^{-1}C_{11}) = 5.253\) and \(\text{tr}((A^{-1}C_{11})^2) = 3.792\). It is important to notice that this algorithm requires knowledge of \(C_{11}\). This requires inversion of a matrix of the order of the number of elements in \(a\). Equations (8) for this example are:

\[
\begin{bmatrix}
\hat{\alpha}_1^2 \\
\hat{\alpha}_2^2 \\
\hat{\alpha}_3^2 \\
\hat{\alpha}_4^2 \\
\hat{\alpha}_5^2 \\
\hat{\alpha}_6^2 \\
\hat{\alpha}_7^2 \\
\end{bmatrix} =
\begin{bmatrix}
(8 - 2 \times 5.253) + 3.792 \\
8 - 5.253 \\
.366 \\
2.905 \\
\end{bmatrix}
\begin{bmatrix}
5.253 - 3.792 \\
8 - 1 - 8 + 5.253 \\
4.713 \\
13.359 \\
\end{bmatrix}^{-1}
\]

The important point to notice is that MIVQUE, in contrast with the Method 3 estimator, makes use of all the data; all eight records and all the correlated structure among them are used to compute estimates of genetic parameters. This property of the estimator can be exploited under the animal model, where each individual's breeding value is predicted regardless of whether they produced offspring or not. With an animal model and a complete matrix of additive genetic relationships, MIVQUE yields unbiased estimates of the base population additive genetic variance, provided that the initial sample of individuals was drawn at random.

However, use of MIVQUE on a sire model that ignores all but half-sib relationships for instance, will not yield unbiased estimates of the base population additive variance because the model does not account for all the correlated structure in the data (Sorensen and Kennedy, 1984b).

An important property of MIVQUE on an animal model is that seems to yield unbiased estimates of the base population additive variance when used on data that have been generated by several cycles of selection and mating (Sorensen and Kennedy, 1984b). This property holds if good prior values of the variances to be estimated are available. However, if the prior is larger than the true value, the estimate is biased downwards and the opposite holds when the prior is smaller than the true value. Monte Carlo results from Sorensen and Kennedy (1984b) suggest that the bias is small; of the order of 10% when the difference between the prior and the true value is of about 40%. More complicated models than the one used in their simulation, with highly unbalanced data may accentuate this difference.

Knowledge of the prior to get unbiased estimates using MIVQUE can be relaxed using iterative MIVQUE or a REML algo-
rithm. Computing costs, however, increase very substantially using this approach, and more understanding is needed on the behavior of different algorithms on different models with respect to their rates of convergence.

Often a reparameterization of a model can lead to considerable savings in computing costs, or can transform a problem from completely unmanageable to manageable. An example is the reduced animal model proposed by Quaas and Pollak (1980). The general idea is that breeding values of later generations, such as offspring breeding values, can be expressed in terms of breeding values of the parental generation plus terms due to Mendelian sampling. In the case of offspring and parent data, for example, this approach reduces the size of the vector of breeding values to the number of individuals in the parental generation. We now illustrate the use of MIVQUE to estimate $\sigma_a^2$ and $\sigma_e^2$ under the reduced animal model.

**MIVQUE Under the Reduced Animal Model.**

We briefly review the properties of the reduced animal model. Consider the following model (one record for each individual):

$$ y = Xb + Ia + \epsilon, \quad (9) $$

where $y$ is a $t \times 1$ vector of records, $b$ is a $p \times 1$ vector of unknown fixed effects, $a$ is a $t \times 1$ vector of additive genetic values of the $t$ individuals, $\epsilon$ is a $t \times 1$ vector of environmental effects, $X$ is a $t \times p$ incidence matrix and $I$ is the identity matrix of order $t$. The first and second moments of the model (ignoring selection) are

$$ E(y) = Xb \quad (10) $$

$$ \text{Var}(y) = V = A\sigma_a^2 + I\sigma_e^2, \quad (11) $$

where all the elements of (11) have been defined previously in (6). We assume for purposes of illustration that the data comprise parents and offspring. There is a total of $N$ parents and $M$ offspring, such that $t = N + M$. Any of the breeding values of the $M$ offspring can be written in terms of the breeding values of its parents. For example if the parents of offspring $j$ are $s$ and $d$, we can write:

$$ y_j = \text{fixed effects} + a_j + \epsilon_j $$

$$ = \text{fixed effects} + 1/2 a_s + 1/2 a_d + \epsilon_j, \quad (12) $$

where $\text{Var}(\epsilon_j) = 1/2 \sigma_a^2 (1 - F) + \sigma_e^2$ and $F$ is the average inbreeding coefficient of $s$ and $d$.

It becomes immediately obvious that by using (12) the number of random effects is reduced from $t$, the total number of individuals, to $N$, the number of parents.

We can formulate the model in matrix notation as follows. Partition $y$ into parental records, $y_p$ and offspring records, $y_o$. Then,

$$ y = \begin{bmatrix} y_p \\ y_o \end{bmatrix} = Xb + \begin{bmatrix} I^T \\ Z \end{bmatrix} a_p + \begin{bmatrix} \epsilon \\ \epsilon^* \end{bmatrix} \quad (13) $$

where, $I$ is the identity of order $N$; $Z$ is a matrix of order $M \times N$ with elements 1/2 or 0 that relate offspring to their parents; $a_p$ is the vector of parental additive genetic values of order $N$; $\epsilon$ is the vector of order $Nx1$ of environmental deviations of the $N$ parents and $\epsilon^*$ is the random error of order $M$ associated with the offspring. We assume that $\text{Var}(a_p) = A_p \sigma_a^2$ and

$$ \text{Var} \begin{bmatrix} \epsilon \\ \epsilon^* \end{bmatrix} = \begin{bmatrix} I & 0 \\ O & K \sigma_a^2/\sigma_e^2 + I \end{bmatrix} \sigma_e^2 = Do \sigma_e^2, \quad (14) $$

where $A_p$ is the $N \times N$ matrix of additive genetic relationships among the $N$ parents, $K$ is a diagonal matrix with its $i$th diagonal element equal to $1/2 (1 - F_i)$, $F_i$ is the average inbreeding of the parents of $i$ and $I$ and $O$ are the identity and a matrix of zeroes respectively.

Let

$$ W = \begin{bmatrix} I \\ Z \end{bmatrix} \quad \text{and} \quad e = \begin{bmatrix} \epsilon \\ \epsilon^* \end{bmatrix} $$

Then (13) is written as:

$$ y = Xb + Wa_p + e. \quad (15) $$

The first two moments of (15) are:

$$ E(y) = Xb \quad (16) $$

$$ \text{Var}(y) = WA_pW' \sigma_a^2 + Do \sigma_e^2. \quad (17) $$

It is easy to show that (17) is equal to (11) and hence (9) and (15) are equivalent models. Quaas and Pollak (1980) show that this reduced animal model is equivalent to absorbing the offspring equations into the parental equations.

We now derive MIVQUE of $\sigma_a^2$ and $\sigma_e^2$ using model (15). Rewrite (17) as follows:

$$ \text{Var}(y) = V = V_1 \sigma_a^2 + V_0 \sigma_e^2, \quad (18) $$
where
\[ V_1 = W_A p W' + \begin{bmatrix} O & O \\ O & K \end{bmatrix} \]  \hspace{1cm} (19)

and
\[ V_0 = I. \] \hspace{1cm} (20)

LaMotte (1970) showed that quadratic forms that yield MIVQUE are:
\[ (y - \hat{X}b)' H^{-1} V_i H^{-1} (y - \hat{X}b), (i=0, 1) \] \hspace{1cm} (21)

In (21), \( \hat{b} \) is a solution to
\[ (X'H^{-1}X) \hat{b} = X'H^{-1}y, \] \hspace{1cm} (22)

where
\[ H^{-1} = R^{-1} - R^{-1} W(W'R^{-1}W + G_p^{-1})^{-1} W'R^{-1}. \] \hspace{1cm} (23)

In (23),
\[ R = \begin{bmatrix} I & O \\ O & \sigma_e^2 \end{bmatrix}, \]
\[ G_p = A_p \sigma_a^2, \]
and \( \gamma = \sigma_e^2 / \sigma_a^2. \)

\( \sigma_e^2 \) and \( \sigma_a^2 \) are priori values of \( \sigma_e^2 \) and of \( \sigma_a^2; \) hence \( H \) is an a priori value of \( V \) (Henderson et al., 1959).

From (20), for \( i=0 \), the LaMotte quadratic is
\[ (y - X\hat{b})' H^{-1} H^{-1} (y - X\hat{b}) = y' P P y, \] \hspace{1cm} (24)

where \( P = H^{-1} [I - X(X'H^{-1}X)X'H^{-1}]. \)

From (19), for \( i = 1 \), the LaMotte quadratic is
\[ (y - X\hat{b})' H^{-1} V_i H^{-1} (y - X\hat{b}) = y' P V_i P y. \] \hspace{1cm} (25)

Hence, MIVQUE of \( \sigma_a^2 \) and of \( \sigma_e^2 \) are obtained by equating (24) and (25) to their expected values. That is:

\[ \begin{bmatrix} \hat{\sigma}_a^2 \\ \hat{\sigma}_e^2 \end{bmatrix} = \begin{bmatrix} \text{tr}(P V_1 P V_1) \\ \text{tr}(P P) \end{bmatrix} ^{-1} \begin{bmatrix} y' P V_1 P y \\ y' P P y \end{bmatrix}. \] \hspace{1cm} (26)

Using the reduced animal model the computational requirement of this estimator is reduced to having to obtain \( H^{-1} \) from (23), which entails inverting a matrix of the order of the number of parents.

This approach can be used to estimate the variance at generation \( t \), with data from generation \( t \) and \( t+1 \). If selection has operated, it is important to include in \( t \) all the records, including those that did not produce any offspring. More efficient algorithms than the one suggested here are needed to use the reduced animal model with large data sets.

We illustrate MIVQUE of \( \sigma_a^2 \) and \( \sigma_e^2 \) with the previous example. As we did in the case of full MIVQUE, we assume that the prior value of \( \sigma_e^2 / \sigma_a^2 \) is 1. In terms of (15), the term \( W_A p \) is:

\[ W_A p = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 \\ 0 & 1/2 & 1/2 \end{bmatrix}. \]

The elements of the diagonal matrix \( R \) are:
\[ 1, 1, 1, 1, 1/2, 1 1/2, 1 1/2, 1 1/2. \]

In terms of computing requirements, the critical step is the inversion of \( [W'R^{-1}W + G_p^{-1}] \) in (23). For this example, this involves the inverse of

\[ W'R^{-1}W + G_p^{-1} = \begin{bmatrix} 2.333 & 0 & .166 \\ 2.333 & .333 & 0 \\ \text{sym.} & 2.5 & 0 \\ 2.166 & 2.5 \end{bmatrix}, \]

where \( G_p \) is the covariance matrix among parents, which is a 4x4 identity matrix in the example. Notice that in contrast with the full MIVQUE, this requires inverting a matrix of the order of the number of parents, as opposed to the number parents plus offspring. The ele-
ments of the first row of $H^{-1}$ are: .567, -.004, .029, .033, -.135, -.133, .008, .008; the diagonal elements of $H^{-1}$ are: .567, .563, .590, .536, .580, .574, .586, .586. The elements of the first row of $P$ and its diagonal elements are, respectively: (.510, -.062, -.019, -.032, -.170, -.174, -.027, -.027) and (.510, .504, .548, .461, .558, .545, .563, .563). Equations (26) and the estimates of $\sigma_a^2$ and $\sigma_e^2$ using the reduced animal model are:

\[
\begin{bmatrix}
\sigma_a^2 \\
\sigma_e^2 \\
\end{bmatrix}
= \begin{bmatrix}
1.2864 & 1.4606 \\
1.4606 & 2.7924 \\
\end{bmatrix}^{-1}
\begin{bmatrix}
4.7132 \\
8.6456 \\
\end{bmatrix}
= \begin{bmatrix}
.366 \\
2.905 \\
\end{bmatrix},
\]

the same estimates obtained under full MIVQUE.

**Estimation of Response to Selection**

In a previous paper, Sorensen and Kennedy (1984a) discussed and compared properties of the least-squares estimator of selection response and an alternative estimator based on mixed model procedures. In this section we clarify and extend some of the results of that paper. We first discuss the traditional least-squares (LS) estimator of response.

When records have been correctly adjusted for fixed effects, then a linear model that can describe a selected line is

\[Y_{ij} = g_i + e_{ij},\]  
(27)

where $g_i$ is the genetic mean of generation $i$ and $e_{ij}$ is the error term. Assume that the vector of records, $y$, comprises parental, $y_0$, and offspring records $y_i$. In matrix form, we have

\[y = Qg + e.\]  
(28)

In (28), $K'$ is a matrix that averages records within generations.

To find the expected value of the LS estimator, we must know the distribution of the vector $y$ and of the breeding values after selection. Assume that the true model is

\[y = \begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mu + \begin{bmatrix} I_0 a_0 \\ I_1 a_1 \end{bmatrix} + \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix},\]  
(30)

where $I$ is a column vector of ones, $\mu$ is the mean before selection and $I$ is the identity matrix and $a_i$ and $\xi_i$ represent additive genetic and environmental values of individuals of the $i$th generation. In (30) we ignore fixed effects for simplicity and assume one record per individual. The variance structure of (30) is:

\[
\text{Var}(y) = V = \begin{bmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{bmatrix} \sigma_a^2 + R \sigma_e^2 = \begin{bmatrix} G_{00} & G_{01} \\ G_{10} & G_{11} \end{bmatrix} + \begin{bmatrix} R_{00} & 0 \\ 0 & R_{11} \end{bmatrix},\]  
(31)

Assume selection was of the form $L'y_0$ and that the deviation of the mean of $L'y_0$ from $I\mu$ is $t$. Then, following Henderson (1975), among selected individuals:

\[
E \begin{bmatrix}
y_0/L'y_0 \\
a_0/L'y_0 \\
\xi_0/L'y_0 \\
\end{bmatrix} = \begin{bmatrix} 1 \mu \\
0 \\
0 \end{bmatrix} + \begin{bmatrix} (G_{00} + R_{00})LH^{-1}t \\
G_{00}LH^{-1}t \\
R_{00}LH^{-1}t \end{bmatrix},\]  
(32)

where $H = \text{Var}(L'y_0)$. The point to notice in (32) is the well-known fact that among the selected group, the mean phenotypic value does not equal the mean breeding value because environmental deviations do not have zero expectation. Sorensen and Kennedy (1984a) inappropriately used (32), which applies to a culling type of selection model, to derive properties of the LS estimator in a...
A model that assumes repeated cycles of selection and mating. After recombination, among the offspring of selected parents, the conditional expectations are:

\[
\begin{bmatrix}
E[y_1/L, y_0] \\
\alpha_1/L, y_0 \\
\epsilon_1/L, y_0
\end{bmatrix} =
\begin{bmatrix}
1 \mu \\
0 \\
0
\end{bmatrix} + \begin{bmatrix}
G_{10}LH^{-1}t \\
0 \\
0
\end{bmatrix}.
\] (33)

The response \(K_1\alpha_1\) has expectation equal to \(K_1G_{10}LH^{-1}t\). If the response is computed as a deviation from the mean of the parental (unselected) generation, \(y_0\), then the expected value of the LS estimator, \(y_1 - y_0\), can be shown to be equal to \(E(K_1\alpha_1)\). The estimator is, of course, also unbiased if \(y_0\) is defined as the phenotypic mean of a contemporaneous control.

We have assumed that records had been adjusted for fixed effects. Using LS procedures this may not always be possible because of confounding between the fixed effect and the breeding value. Further, if for example, due to the correlated structure of the data, there are different amounts of information on the offspring of selected families (i.e., unequal family sizes) the LS estimator ignores this fact and each individual contributes equally to the generation mean. This results in sampling variances of response to selection that are larger than those of an estimator that makes use of this information.

If the response at time \(t\), \(R_t\), is estimated as the deviation of the mean phenotypic value of the selected and an independent control line sampled from the same base population, then \(R_t = \hat{\gamma}_t^s - \hat{\gamma}_t^c\), where superscripts \(s\) and \(c\) represent selected and control, respectively. The variance of estimated response is

\[
\text{Var}(\hat{R}_t) = \text{Var}(\hat{\gamma}_t^s) + \text{Var}(\hat{\gamma}_t^c). \quad (34)
\]

In the control line, the variance of the vector of generation means is (Sorensen and Kennedy, 1983):

\[
\text{Var}(\hat{\gamma}^c) = \begin{bmatrix}
\tilde{a}_0 & \tilde{a}_0 & \tilde{a}_0 & \cdots & \tilde{a}_0 \\
\tilde{a}_0 & \tilde{a}_1 & \tilde{a}_1 & \cdots & \tilde{a}_1 \\
\tilde{a}_0 & \tilde{a}_1 & \tilde{a}_2 & \cdots & \tilde{a}_2 \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
\tilde{a}_0 & \tilde{a}_1 & \tilde{a}_2 & \cdots & \tilde{a}_t
\end{bmatrix} \sigma_a^2 + 10\sigma_e^2 / N,
\]

where \(\tilde{a}_i\) is the average additive relationship among the \(N\) breeding values of generation \(i\), and it represents drift variance that accumulates each generation (Hill, 1971) such that \(\tilde{a}_0 < \tilde{a}_1 < \tilde{a}_2 < \ldots < \tilde{a}_t\). At generation \(t\),

\[
\text{Var}(\hat{\gamma}_t^c) = (\tilde{a}_t h^2 + (1-h^2)/N)\sigma_a^2 \quad (36)
\]

where \(\sigma_a^2\) is the phenotypic variance in the base population. If we ignore the complications of selection on the variance structure of the data (see Hill, 1977, 1979 for a discussion) and assume that selected and control lines are kept with equal effective number of parents each generation, then

\[
\text{Var}(\hat{R}_t) = 2\sigma_a^2(\tilde{a}_t h^2 + (1-h^2)/N), \quad (37)
\]

where \(\sigma_a^2\) is the phenotypic variance in the base population.

We illustrate some of these points with the same data as before:

\[
y = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{g}_0 \\ \hat{g}_1 \end{bmatrix} + e.
\]

The least-squares estimator is:

\[
\hat{g} = \begin{bmatrix} \hat{g}_0 \\ \hat{g}_1 \end{bmatrix} = \begin{bmatrix} 1/4 & 0 \\ 0 & 1/4 \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \begin{bmatrix} \hat{y}_0 \\ \hat{y}_1 \end{bmatrix} = \begin{bmatrix} 3.5 \\ 3.5 \end{bmatrix}.
\]

The variance of \(\hat{g}\) is:

\[
\text{Var}(\hat{g}) = (Q'Q)^{-1}Q' \text{Var}(y)Q(Q'Q)^{-1} = \begin{bmatrix} \tilde{a}_0 & \tilde{a}_0 \\ \tilde{a}_0 & \tilde{a}_1 \end{bmatrix} \sigma_a^2 + (Q'Q)^{-1} \sigma_e^2
\]

where \(\tilde{a}_i\) is the average relationship among individuals of generation \(i\), including relationship to self (Sorensen and Kennedy, 1983). As we did before, we assume that \(\sigma_a^2 = \sigma_e^2 = 1\). Hence,
The least-squares estimate of response and its standard error is \(.00 \pm .81\). In this simple example we assumed that no control population was necessary to obtain an unbiased estimate of response. If a control is necessary to correct for environmental trend, the variance of the estimate of response is increased relative to its value without a control. For example, using (37), \(\text{Var}(\hat{R}_t) = 1.3125\) and the standard error is 1.15 units.

An alternative way of estimating response to selection is to use a mixed model approach. Henderson (1975) has shown that when selection involves culling on the basis of past performance then, under certain conditions, the mixed model equations ignoring selection lead to best linear unbiased estimators (BLUE) of estimable functions of fixed effects and best linear unbiased predictors (BLUP) of the random effects of the model. These conditions are: (i) the model is the correct one; (ii) selection is on a linear function of the records, (iii) the ratios of the variances of the random effects prior to selection are known, (iv) the random effects before selection are multivariate normally distributed and (v) selection is invariant to the fixed effects in the model, namely, \(L'X = 0\). Henderson (1980) has recently given examples of types of selection that yield \(L'X = 0\). Henderson (1980) has shown that if the above conditions hold, then \(\hat{b}\) and \(\hat{a}\) obtained from (38) are BLUE of \(b\), assuming \(b\) is estimable, and BLUP of \(a\). It can readily be shown that if these conditions hold, the estimator of response \(K'\hat{a}\) is unbiased, i.e., \(E(K'\hat{a}) = E(K'a)\), and it is therefore arguable whether a control is at all needed since the phenotypic trend can be partitioned into its genetic and environmental components. The conditions which make this possible are rather restrictive, particularly (i) and (iii). We discuss this point further, below.

The sampling variance of the estimate of genetic means is computed as follows. For simplicity, assume the model is

\[
y_t = a_t + e_t. \tag{39}
\]

The vector of the estimates of genetic means is \(K'\hat{a}\), where in this completely random model \(\hat{a}\) is the best linear predictor of \(a\) in (39). Then

\[
\text{Var}(\hat{a}) = A\sigma_a^2 - C\sigma_e^2, \tag{40}
\]

where \(C = (I + A^{-1} \gamma)^{-1}\). Notice that \(C = \text{Var}(\hat{a} - a)\), the prediction error variance. Then, the variance of the estimate of genetic means in generation \(t\) is:

\[
\text{Var}(K_t'\hat{a}) = (\bar{a}_t h^2 - c_t (1-h^2)) \sigma_e^2. \tag{41}
\]

Because \(c_t > 0\), the variance in (41) is smaller than the variance of the LS estimator of response in (36). As in the case of the LS estimator, with a very large number of observations, the variance of the estimator reduces to the variance due to drift, \(a_t h^2 \sigma^2 = 2F_t h^2 \sigma^2\). Notice that in (41), the variance increases with increasing information. This is a consequence of the result that in BLUP and in selection index, the prediction error variance decreases and the variance of the predictor increases with increasing information. Of course, the drift term in (41) increases with a smaller number of parents.

So far it has been assumed that the heritability in the base population is known and used in the mixed model equations (38) to compute response. It is arguable and more than of academic interest whether one should label this an estimator or a predictor of response. Strictly, a prediction of response is data independent. The mixed model approach is data-dependent and can probably be best
characterized as a form of Bayesian estimation. However, the use of $A^{-1}$ has the effect of comparing an animal's record with an index based on the animal's relatives. If the observed value is different from the index, the animal's predicted breeding value is regressed towards the observed value (Schaeffer, 1982). Because of this property, estimated response will lie between the true response and the predicted one. There are two consequences of this result. First, the computed response tends to be smoother than the true response (or than that obtained by LS) because it tends to be less affected by true sampling deviations, i.e., formula (41). Second, if observed response differs from that predicted not because of drift but because of the model on which the predicted value is based does not hold, then the computed response will be biased. We discuss this point below.

What can be done if the initial heritability is not known? At least two approaches can be followed. One is to use a prior based on information from the literature; the other is to obtain an estimate from the data. Sorensen and Kennedy (1984a) tested the use of a wrong prior in their simulation to compute response using mixed model procedures. When the prior is regarded as a constant, Henderson (1975) shows that predicted breeding values are biased, the bias being proportional to $(a^2 - a_f^2)$, where $a_f^2$ is the prior additive variance. In their simulation results, Sorensen and Kennedy (1984a) used prior heritabilities of .3 and .7, when in fact true heritability was .5. The true response after two cycles of selection was of 2.6 units. Using the prior of .3, the response was overestimated by .21 units and underestimated by the same amount using the prior of .7. Hence a difference between the prior and the parameter of 40% led to a bias of about 8%.

Alternatively, one can obtain an estimate of the base population heritability using a MIVQUE or REML estimator. The estimate can then be used in the mixed model equations to compute response. Notice that in this approach, the prior is not a constant but a random variable. If the estimator is unbiased, it seems intuitively that, over a conceptually large number of lines, the expected estimate of response should equal the true response, assuming the model is correct. Kackar and Harville (1981) have shown this to be so for the no-selection case; similar results are not known to us for the selection situation.

<table>
<thead>
<tr>
<th>Generation</th>
<th>True genetic mean</th>
<th>Estimated genetic mean</th>
<th>BLUP</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Mixed model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>0.94</td>
<td>1.00</td>
<td>0.04</td>
<td>1.23</td>
<td>1.46</td>
<td>1.29</td>
</tr>
<tr>
<td>2</td>
<td>1.80</td>
<td>1.89</td>
<td>1.91</td>
<td>0.03</td>
<td>2.50</td>
<td>2.45</td>
<td>2.49</td>
</tr>
<tr>
<td>3</td>
<td>1.23</td>
<td>1.23</td>
<td>1.34</td>
<td>0.34</td>
<td>2.46</td>
<td>2.46</td>
<td>2.46</td>
</tr>
</tbody>
</table>

*Based on 200 Monte Carlo replicates. Within each replicate, five out of 20 are selected each generation on phenotypic performance. Selection is only on males and initial heritability is 50%. 

TABLE 1. MEANS AND EMPirical STANDARD DEVIATIONS OF TRUE GENETIC MEAN AND ESTIMATED GENETIC MEAN BY LEAST-SQUARES, BEST LINEAR UNBIASED PREDICTION WITH INITIAL HERITABILITY KNOWN (BLUP) AND MIXED MODEL METHODS USING A MIVQUE ESTIMATE OF INITIAL HERITABILITY OBTAINED FROM THE DATA.
Table 1 shows results from Monte Carlo simulation where the true response and its variance is compared with the following estimates: (i) a LS estimate, (ii) a mixed model (BLUP) estimate when initial heritability is known and (iii) a mixed model estimate which uses as a prior a MIVQUE estimate of heritability from the same data. Details of the simulation techniques are as in Sorensen and Kennedy (1984a). The points to notice are that all estimators seem to be unbiased; that (ii) has the smallest variance, smaller than the variance of true response; that the variance of (iii) is larger than the true variance and smaller than that of (i). The variances of both the LS and of the mixed model estimators also were computed from (36) and from (41). Neither expression holds exactly under selection or when heritabilities are not known but both gave results in good agreement with the empirical variances obtained from the simulation (table 2).

As mentioned earlier, an important assumption of the mixed model approach is that the model describing the data is correct. If a model such as (5) is used, the experimenter assumes that a strictly additive model holds for the trait selected for, i.e., no dominance, maternal effects, etc. For unbiasedness in the selection case, the mixed model requires that the distribution of \( a \) and \( \gamma \) be multivariate normal. The genetic implication of this is that the trait is affected by a very large number of additive loci in which case, selection followed by random mating causes only small departures from normality (Bulmer, 1971). With finite numbers of loci that move towards fixation, this assumption will not hold. Further, the mixed model also requires that selection is on a linear function of the records and that the selection differential is not contributed by other unknown traits correlated with the selected trait, such as natural selection. Thus, there is an implied circularity in the mixed model approach; experiments are often conducted to check whether observed responses agree with those predicted under a given model. However, for the mixed model to yield unbiased estimates of response, the model has to be the correct one. Thus the price to pay for a smaller variance of the estimator may be a bias in the estimate of response. The magnitude of the bias will depend on the degree to which the assumptions of the operational model are violated. These are likely to be more serious in long-term selection experiments that span several generations. However in short-term selection experiments with domestic animals, the mean square error of the mixed model estimator, which is the sum of its variance plus the square of the bias, may still be considerably smaller than that of the least-squares estimator. This may be particularly relevant when the limitations of space and resources may force the experimenter to choose between either allocating facilities to both a selected and a control line, or devoting them all to a selected line with twice the effective population size. The possibilities of using a mixed model approach to analyse response may have a bearing on the design of the experiment. It seems to us important, to carry out critical simulation studies with a variety of genetic models, to test the robustness of the mixed model estimator.

We illustrate the use of the mixed model approach using the data from the previous examples. The mixed model equations and their

---

**Table 2. Observed and Predicted Standard Deviations of Least-Squares and Mixed Model Estimators of Genetic Means**

<table>
<thead>
<tr>
<th>Generation</th>
<th>Least-squares Observed</th>
<th>Least-squares Predicted</th>
<th>Mixed model Observed</th>
<th>Mixed model Predicted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.01</td>
<td>.98</td>
<td>.94</td>
<td>.89</td>
</tr>
<tr>
<td>2</td>
<td>1.18</td>
<td>1.20</td>
<td>1.09</td>
<td>1.08</td>
</tr>
<tr>
<td>3</td>
<td>1.34</td>
<td>1.38</td>
<td>1.29</td>
<td>1.25</td>
</tr>
</tbody>
</table>

*Based on 200 Monte Carlo replicates.

*From formula (36) with \( \sigma_a^2 \) and \( \sigma_e^2 \) known.

*From formula (41) with MIVQUE estimates of \( \sigma_a^2 \) and \( \sigma_e^2 \).
solution \( (\hat{\alpha}) \) for this example were obtained previously. The estimate of response is obtained from \( K'\hat{\alpha} \):

\[
K'\hat{\alpha} = \begin{bmatrix}
\frac{1}{4} & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\
\end{bmatrix} \hat{\alpha} = \begin{bmatrix} 0 \\
-1.135 \end{bmatrix}.
\]

Notice that the average breeding value of individuals that comprise the base is forced to zero. The variance of the computed response is:

\[
\operatorname{Var}(K'\hat{\alpha}) = K' \operatorname{Var}(\hat{\alpha}) K, \text{ where } \operatorname{Var}(\hat{\alpha}) = A\sigma^2_a - C\sigma^2_e.
\]

Hence, \( \operatorname{Var}(K'\hat{\alpha}) = \begin{bmatrix} 0 & 0 \\
0 & 0.0475 \end{bmatrix} \).

The mixed model estimate of response and its standard error is \(-1.14 \pm 0.22\). Notice that the variance using this approach is considerably smaller than the variance of the least-squares estimator.

**Conclusions**

In this paper we have illustrated the use of mixed model techniques and have discussed some of the properties of the estimators and limitations. Mixed model analysis of data is routinely undertaken to analyze field records where the data are highly unbalanced and there are large numbers of fixed effects and often partial confounding. Selection from carefully designed experiments, however, particularly with laboratory animals such as mice and Drosophila generate data that can be analyzed with simple traditional estimators, and a mixed model approach is not likely to contribute further insight.

Many selection experiments, however, are carried out with domestic animals where generations overlap and the correct specification of the model must include fixed effects such as sex, age of dam, year effects, etc. In these cases, a mixed model approach is likely to show considerable advantages over the least-squares estimator. With fixed effects and several generations, simple estimators of heritability based on offspring-parent regression are of little use. The regression model can only correct for fixed effects on the offspring records but not on the parents. On the other hand, a mixed-model type estimator such as MIVQUE can easily accommodate fixed effects and uses all the data to obtain estimates of initial genetic variances. In the years to come, the computational demands of such estimators are not likely to be a limiting factor in their use.

We have discussed the use of a mixed model approach to estimate response to selection, and have described the conditions under which this estimator adequately partitions phenotypic trend into its genetic and environmental components. We do not wish this to imply that selection experiments should be designed without the use of contemporaneous controls.

In an ideal situation where facilities are available, selection experiments should be designed not only with unselected controls, but should also be replicated. Without adequate replication, interpretation of results is likely to be ambiguous, particularly in the case of correlated response, because often each line has its own specific behavior. However, if resources are severely limited, the experimenter may well wish to consider the option of eliminating the control line and to devote the facilities to selection lines, and use a mixed model approach to analyze the data.

**Literature Cited**


Thompson, R. 1977. The estimation of heritability with unbalanced data. II. Data available on more than two generations. Biometrics 33:497.