Choosing Mating Designs to Efficiently Estimate Genetic Variance Components for Trees

1. Sampling Errors of Standard Analysis of Variance Estimators

By G. Namkong and J. H. Rorberd

(Received February 1974)

The choice of a mating design for estimating genetic variances in forestry should be dictated by the likelihood of achieving the estimation objectives within the usual restrictions of time, space, cost, and biological limitations. Poor designs can cause poor estimates or failure to detect even very high genetic variances. Since experiments with trees are relatively expensive and time consuming, foresters try to estimate several parameters in the same experiment. This paper compares the relative efficiencies of standard, balanced, mating designs for estimating additive and dominance genetic variances. The main variables considered are mating design, total number of crosses, total number of parents, number of half-sib versus full-sib families, level of the genetic variances, and the use of sub-blocking. These are special cases of the general problem of estimating variance components (for a good review see Seable, 1971) expressed in a particular useful way for plant breeders.

The need for precision is often dependent on the relative size of the parameter which is being estimated. When the genetic variance is twice the nongenetic variance, a precise estimate of the genetic variance may not be necessary. When genetic variance is low for a trait however, interest in the case, estimation errors should be evaluated over all levels of heritability. Clearly, many other value functions can be used to evaluate designs in terms of the heritability. Since errors of estimation are also dependent on heritability, we examine efficiency over a continuous range of heritabilities.

Mating Designs

Consideration is restricted to designs using only one common generation of parental materials originating as a random sample from a single random mating population. The designs are the nested (AB) factorial (AB), and diallel (AA) as designated by Cockerham (1963). All these designs are well known, and are described elsewhere (see Cockerham, 1963; Gardner, 1963). All permit estimates of full-sib, half-sib, and error variance components which can be interpreted in terms of genetic variances of the sampled population. Two useful variants of the diallel design are investigated here; the partial diallels of Hinkelman and Stern (1960), and Kemphorine and Curnow (1961), and the disconnected diallels of Braaten (1965). For a given number of crosses, the partial diallels are always more efficient than the nested design (AB) for both additive and dominance genetic variance components (Kemphorine and Curnow, 1961). Similarly, the disconnected diallels are more efficient than the factorial (AB) for estimating the additive genetic variance component at some heritabilities for the same number of crosses. These mating designs are schematically drawn in Figure 1. Their analyses of variance are listed in Table 1 with their derived estimates of the additive genetic variance, $\sigma^2_{A}$, and the dominance genetic variance, $\sigma^2_{D}$, under the simplifying assumptions of no epistasis, no genotype by environment interaction, regular diploid inheritance, and no disequilibria of any kind.

We shall compare ratios of the sampling variances of the estimates of the additive genetic variance, $V(\delta^2_A)$ to the $\delta^2_A$ itself. Similar ratios for the dominance genetic variance are $V(\delta^2_D)/\delta^2_D$. Designs and various allocations of males and females are compared first for an equivalent number of crosses, then for an equivalent number of parents. If the work of making the crosses or conducting the field experiment is the limiting factor, then comparison should be based on equivalent numbers of crosses. If a set of parents is previously established, then the designs should be compared with equivalent numbers of parents. In such cases, there is some increased cost with increased numbers of crosses, but the additional costs are likely to be a non-linear function of the numbers. We do not investigate these costs but merely indicate that cost analyses should be applied to our functions.

Randomized complete blocks within gross environmental classes are assumed in this discussion and as many environments as possible will be sampled. Hence,
the replication mean square is a pooled mean square of replications within several environments. For the general equations, the replication number is a variable (r) coefficient, but for the illustrations we arbitrarily assume 10 replications. Fewer replications increase the errors of estimation.

We also generally assume that there are sufficient trees per family plot within a replication for realistic and well-distributed plot means. In the general equations, the number of trees per plot is a variable (n) but in the illustrations eight is arbitrarily used. More trees per plot would clearly reduce the plot error but would reduce the number of trees available for sampling many environments. Consider for example, that the variance among two individual trees from the same family, randomly planted within a replicate, contains an environmental component, \( \sigma_{e}^{2} \), and a genetic variance among individuals within families, \( \sigma_{g}^{2} \). Then if the trees are grown in plots of size n, the plot variance is \( \sigma_{p}^{2} = \sigma_{g}^{2}/n + \sigma_{e}^{2}/n \) or, using H. F. Smith's (1938) empiric relationship for the environmental error component, \( \sigma_{p}^{2} = \sigma_{g}^{2}/n + \sigma_{e}^{2}/n^{b} \), where b is Smith's (1938) soil heterogeneity coefficient. Then the remaining variance among individuals within plots, \( \sigma_{i}^{2} \), is \( \sigma_{i}^{2} = \sigma_{g}^{2} - \sigma_{p}^{2} = \sigma_{g}^{2}/n[(n-1)/n] + \sigma_{e}^{2}/n^{b}(n^{b-1}/n^{b}) \)

The plot error mean square on a plot mean basis is:

\[
\text{Error M.S.} = \sigma_{e}^{2}/n \cdot \sigma_{p}^{2} = \sigma_{g}^{2}/(2n - 1)/n^{b} + \sigma_{e}^{2}/(2n^{b-1}/n^{b})
\]

In Table 1, this is equal to the \( \sigma^{2} \) for the expected value of the error mean square.

Furthermore, if environmental heterogeneity is such that large replication blocks are not feasible, then sub-blocking to reduce "within replication" error variation may be desirable. The greater the heterogeneity, the more sub-blocks would be desirable. Using H. F. Smith's (1938) empiric equation for relating soil heterogeneity to block size, the \( \sigma_{p}^{2} \) factor in the above equation for the plot error mean square should be multiplied by a coefficient, reducing \( \sigma_{p}^{2} \) according to the soil heterogeneity coefficient, b, the number of sub-blocks, p, and the number of family plots per sub-block, c. The new \( \sigma_{p}^{2} \) say \( \sigma_{p}^{2} \), is:

\[
\sigma_{i}^{2} = \sigma_{g}^{2}/((pc - 1) - cot - c/p - 1/p)(pc - 1)/cot - 1/p(c - 1)) \]

(SCHUTZ and COCKENHAM, 1962). The equations for the errors are all

---

**Figure 1. Submatic Diagram of Four Mating Designs**

- **Females**
  - ABCDEFGHIJL
  - A: X
  - B: X
  - C: X
  - D: X
  - E: X
  - F: X
  - G: X
  - H: X
  - I: X
  - J: X
  - K: X
  - L: X

  - Do female serves as a male
  - Do male serves as a female
  - Do female serves more than one male
  - Each male serves the same number of different females
  - Reciprocal crosses possible (0’s)
  - Sub-blocking different sets may be accomplished by blocking different male groups

- **Females**
  - ABCDEFGHIJL
  - A: X
  - B: X
  - C: X
  - D: X
  - E: X
  - F: X
  - G: X
  - H: X
  - I: X
  - J: X
  - K: X
  - L: X

  - Do female serves as a male
  - Do male serves as a female
  - Each male serves the same set of males
  - Each male serves the same set of females
  - Reciprocal crosses possible (0’s)
  - Sub-blocking sets may be accomplished by blocking male or female groups or by blocking separate male and female factorials

- **Females**
  - ABCDEFGHIJL
  - A: X
  - B: X
  - C: X
  - D: X
  - E: X
  - F: X
  - G: X
  - H: X
  - I: X
  - J: X
  - K: X
  - L: X

  - Males and females at both rates
  - Overlapping crossing among parents
  - Crosses define relationships
  - Reciprocal crosses possible
  - Sub-blocking sets may be accomplished by blocking chains into groups

- **Females**
  - ABCDEFGHIJL
  - A: X
  - B: X
  - C: X
  - D: X
  - E: X
  - F: X
  - G: X
  - H: X
  - I: X
  - J: X
  - K: X
  - L: X

  - Same as partial diallel but separated into groups of small partial diallels
  - Reciprocal crosses possible (0’s)
  - Sub-blocking sets may be accomplished by blocking separate diallels
Table 1. Analyses of Variance of Four Mating Designs and Their Derived Estimates of the Genetic Variance.

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>df</th>
<th>Expected Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>r-1</td>
<td></td>
</tr>
<tr>
<td>Sub-blocks</td>
<td>g-1</td>
<td></td>
</tr>
<tr>
<td>Replications x Sub-blocks</td>
<td>(r-1)(g-1)</td>
<td></td>
</tr>
<tr>
<td>Males (MS)</td>
<td>m-p</td>
<td></td>
</tr>
<tr>
<td>Females (MS/M)</td>
<td>m-1</td>
<td></td>
</tr>
<tr>
<td>Plan Mean Error (MSE)</td>
<td>m(f-1)</td>
<td></td>
</tr>
</tbody>
</table>

\[
\sum s^2 = \sum s^2_i + \sum s^2_{i-1} = \sum (MSE-MSE)/r = \sum s^2_{i-1}
\]

\[
\sum s^2_i = \sum s^2_{i-1} + \sum (MSR-MS)/r = \sum s^2_{i-1} + \sum s^2_{i-2}
\]

where \( r \) = number of replicates
\( g \) = number of sub-blocks
\( m \) = number of male groups in entire experiment
\( f \) = number of females per male group

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>df</th>
<th>Expected Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>r-1</td>
<td></td>
</tr>
<tr>
<td>Sub-blocks</td>
<td>g-1</td>
<td></td>
</tr>
<tr>
<td>Males (MS)</td>
<td>m-p</td>
<td></td>
</tr>
<tr>
<td>Females (MS/M)</td>
<td>m-1</td>
<td></td>
</tr>
<tr>
<td>(Males, Females Pooled) (MS/M)</td>
<td>m(1-p)</td>
<td></td>
</tr>
<tr>
<td>Males x Females (MSSF)</td>
<td>m(p-1)(f-1)/2</td>
<td></td>
</tr>
<tr>
<td>Error (MSE)</td>
<td>m(f-1)(r-1)/2</td>
<td></td>
</tr>
</tbody>
</table>

\[
\sum s^2 = \sum s^2_i + \sum s^2_{i-1} = \sum (MSR-MS)/r = \sum s^2_{i-1} + \sum s^2_{i-2}
\]

\[
\sum s^2_i = \sum s^2_{i-1} + \sum (MSR-MS)/r = \sum s^2_{i-1} + \sum s^2_{i-2}
\]

where \( r \) = number of replicates
\( g \) = number of sub-blocks
\( m \) = number of males in entire experiment
\( f \) = number of females in entire experiment

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>df</th>
<th>Expected Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>r-1</td>
<td></td>
</tr>
<tr>
<td>Sub-blocks</td>
<td>g-1</td>
<td></td>
</tr>
<tr>
<td>General Combining Ability (GCA)</td>
<td>(m-1)/2</td>
<td></td>
</tr>
<tr>
<td>Specific Combining Ability (SCA)</td>
<td>(m-1)(f-1)/2</td>
<td></td>
</tr>
<tr>
<td>Error (MSE)</td>
<td>m(f-1)(g-1)/2</td>
<td></td>
</tr>
</tbody>
</table>

\[
\sum s^2 = \sum s^2_i + \sum s^2_{i-1} = \sum (MSR-MS)/r = \sum s^2_{i-1} + \sum s^2_{i-2}
\]

\[
\sum s^2_i = \sum s^2_{i-1} + \sum (MSR-MS)/r = \sum s^2_{i-1} + \sum s^2_{i-2}
\]

\( s \) = number of parents in entire experiment
\( g \) = number of crosses per parent

<table>
<thead>
<tr>
<th>Source of Variance</th>
<th>df</th>
<th>Expected Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>r-1</td>
<td></td>
</tr>
<tr>
<td>Sub-blocks</td>
<td>g-1</td>
<td></td>
</tr>
<tr>
<td>General Combining Ability (GCA)</td>
<td>(m-1)/2</td>
<td></td>
</tr>
<tr>
<td>Specific Combining Ability (SCA)</td>
<td>(m-1)(f-1)/2</td>
<td></td>
</tr>
<tr>
<td>Error (MSE)</td>
<td>m(f-1)(g-1)/2</td>
<td></td>
</tr>
</tbody>
</table>

\[
\sum s^2 = \sum s^2_i + \sum s^2_{i-1} = \sum (MSR-MS)/r = \sum s^2_{i-1} + \sum s^2_{i-2}
\]

\[
\sum s^2_i = \sum s^2_{i-1} + \sum (MSR-MS)/r = \sum s^2_{i-1} + \sum s^2_{i-2}
\]

\( s \) = number of parents in entire experiment
\( g \) = number of crosses per parent

derived in terms of a general variable \( \sigma^2 \) so that for any level of environmental variance the expected errors of estimation can be derived. The illustrations use \( \sigma^2 = 1\), and it will be shown in the error equations that \( \sigma^2 \) is a simple coefficient of all errors and, therefore, that comparisons among designs for any given \( \sigma^2 \) will be a simple multiple of the comparisons illustrated in this paper.

Estimation of Sampling Variances

From the mean squares of Table 1, the estimators of \( \sigma^2 \) and \( \sigma^2 \) can be seen to be simple linear functions of the mean squares and various design parameters like \( m, r, \) etc. If the plot means are normally distributed, their sums of squares and the partitions in balanced experiments are \( \chi^2 \) variables which are multiples of \( \chi^2 \) variables. The variance of a \( \chi^2 \) variate is \( 2 \cdot (d.f.) \) and its variance when multiplied by a coefficient (MS) is \( (MS)^2 \cdot 2 \cdot (d.f.) \). We estimate the variance by using estimated MS's. The sum of squares (SS) is a \( \chi^2 \) variate with a multiplicative equal to the expected mean square. Therefore, the sampling variance of a SS is \( (MS)^2 \cdot 2 \cdot (d.f.) \) and for a mean square is

\[
V(SS/d.f.) = V(SS) \cdot d.f. = (MS)^2 \cdot 2 \cdot d.f.
\]

For a linear function of mean squares with coefficients, say \( k_0 \), the variance of the linear function is \( 2k_0^2 \cdot (MS)^2 \cdot d.f. \) where summation is taken over all mean squares in the function and there is no covariance between the mean squares. When the component estimated is \( \sigma^2 \), the sampling variance of \( \sigma^2 \) is 16 times the component's sampling variance. The sampling variances for \( \sigma^2 \) and \( \sigma^2 \) for the designs examined are listed in Appendix Table A1.

In the factorial (AB) mating design, the male and female mean squares are pooled and since their EMS are not generally exactly alike, the resulting sum of squares is not a \( \chi^2 \) type of variate. However, the variance of the pooled sum of squares is the variance of a weighted sum of the two original male and female sum of squares which are \( \chi^2 \) type variates. Therefore, the variance of the sum is a sum of these variances: \( 2(m + f - 2p^2)p^2 / (m + f - 2p)^3 \).

Similarly, the mean squares from the diakels are not simple \( \chi^2 \) type variates and their variances have to be derived as a sum of variances. They are given by Braaten (1965).

Since the above variances are linear functions of mean squares and the mean squares are linear functions of the error and genetic variance components, the sampling variances can be translated in terms of the basic variance components themselves. Furthermore, since designs can be evaluated most conveniently on the basis of comparable ranges of heritability, it would be suitable to further translate the variances into functions of heritability. Then it would be very easy to examine relative efficiencies of allocations, designs, numbers of crosses, etc., for any given heritability level or for any range of interest. Geneticists can then determine which sets of experimental allocations have sampling errors small enough to give reasonable chances of detecting genetic variates at any desired heritability levels. As a matter of convenience, the sampling errors are hereafter taken as \( V(\sigma^2) / \sigma^2 \) and \( V(\sigma^2) / \sigma^2 \). The sampling error equations are listed in Appendix Table 2 in terms of the experimental and mating design parameters and heritability.
and the ratio of $\sigma_2^2$ to $\sigma_3^2 + \sigma_4^2$, which we call the dominance ratio, $I$:

$$I = \frac{\sigma_2^2}{\sigma_3^2 + \sigma_4^2}$$

While these functions are rather cumbersome, they can be easily evaluated for variations in any of the parameters.

**Design Comparisons**

The application of the developed formulae to all contrasts of interest to tree breeders would require cumbersome numbers of charts. In this paper, we display the effects of a few variations, such as allocations among designs, as examples of how the formulae in terms of $H$ and $I$ can be useful.

The coefficient of variation (CV) may also be useful in evaluation. This CV is the standard deviation of estimated $\sigma_1^2$ divided by $\sigma_1^2$; $\sqrt{V(\delta_1^2)/\sigma_1^2}$. For any level of the genetic variance or heritability, designs which give lower variances than would be needed for a CV of $\frac{1}{2}$, for example, may be considered good. In such cases, $\sigma_1^2$ will be larger than twice its standard error of estimate. For example, if $\sigma_2^2 = 1$, $\sigma_1^2 = 0$, and $\sigma_1^2 = \frac{1}{2}$, then $H = 0.1$ and a design with expected standard error of $\pm 0.05$ or less for estimating $\sigma_1^2$, would be satisfactory. This criterion may be appropriate if the geneticist is interested in detecting genetic variances whatever the size of the $\sigma_1^2$, or $\sigma_1^2$. On the other hand, a geneticist may not be interested in low heritabilities and would use the CV criterion only at moderate to high heritabilities. At low heritabilities, a more constant error rate, not dependent on $\sigma_2^2$ itself, may be more appropriate. For example, using the same figures of $\sigma_2^2 = 1$ and $\sigma_3^2 = 0$, but $\sigma_1^2 = 0.05$ and $H = 0.048$, an error of $\pm 0.06$ may still be acceptable. We suggest that a reasonable criterion in such an evaluation would be to first determine the minimum heritability of interest. Below that point the sampling error may be allowed to be constant but above that point the CV criterion may be applied. In the following figures the CV = $\frac{1}{2}$ curve is drawn for comparison, and the constant variance curve below $H = 0.1$ and $H = 0.05$ is also drawn.

**General Features of the Curves**

The general shape of the curves when drawn on semi-log paper is similar to a lower case Greek letter $\nu$. The left arms of the curves are nearly vertical near heritabilities of zero, especially in the range of $H$ of 0 to 0.5 reflecting the very small levels of $\sigma_1^2$ in the denominator of $V(\delta_1^2)/\sigma_1^2$. At higher $H$ levels the variance increases faster than $\sigma_1^2$ itself and as $\sigma_1^2$ increases very rapidly as $H$ approaches 1, the function again assumes an almost vertical form.

The level of the entire $\nu$-shaped curve changes with respect to the design parameters as does the slope of the right-hand portion and the point at which the left-hand stem joins the ascending right-hand part.

For any constant CV, say $c$, the $c = \sqrt{V(\delta_1^2)/\sigma_1^2}$, then $c' = V(\delta_1^2)/c\sigma_1^2$, and $c'\omega_1^2 = V(\delta_1^2)/\omega_1^2$. Therefore, for any $c$ value the variable scaled on the ordinate axes of the figures, $V(\delta_1^2)/\omega_1^2$, changes as a multiple of $\omega_1^2$. The ordinate value is zero at $\omega_1^2 = 0$ and therefore when $H = 0$, rises as $H$ increases, and approaches infinity as $\omega_1^2$ approaches infinity, which implies that $H$ approaches one. Such a curve must lie below the $\nu$-shaped curve for any given design in some region of $H$, and hence, no design could be completely satisfactory. On the other hand, a curve of $V(\delta_1^2)/\omega_1^2$ for any constant variance would start very high at low values of $H$ and continue to decline as $H$ approaches one. Using our suggested criterion of a constant variance at low $H$, and a constant CV at moderate to high levels of $H$, results in another $\nu$-shaped curve of critical values. Design curves below the criterion curve for all $H$ values are then possible to define.

![Figure 2. Efficiency of allocations ($V(\delta_1^2)/\sigma_1^2$) for estimating $\sigma_1^2$, when number of crosses ($c$) = 100.](image)

<table>
<thead>
<tr>
<th>Design</th>
<th>Number of Males</th>
<th>Number of Parents</th>
<th>Number of Crosses per Parent</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Nested Design</td>
<td>$M$</td>
<td>$Q$</td>
<td>$S$</td>
</tr>
<tr>
<td>B. Factorial</td>
<td>$F$/$M$</td>
<td>$P$</td>
<td>$C$</td>
</tr>
<tr>
<td>C. Diallel</td>
<td>$P$/$Q$</td>
<td>$M$</td>
<td>$F$</td>
</tr>
</tbody>
</table>
Allocating Degrees of Freedom for Full and Half-Sib Families

Any design permits permutations of many full-sib families per half-sib family or many half-sibs with few full-sib families in each. For estimating $\sigma_A^2$ in the nested design (A/B), the allocations with maximum degrees of freedom for half-sibs are generally superior except at low heritabilities (Fig. 2A). In those cases, the better precision in estimating the error mean squares is a dominating factor and changes the relative efficiency of designs. However, for a large range in heritability, the allocations with many half-sib groups are required to give the investigator a reasonable chance of detecting the genetic variance even when using 100 crosses in experiments like these. A reasonable allocation which would be good for low and high heritabilities would include up to six female full-sib families per male half-sib family.

In the nested designs, the relative efficiencies of different male:female ratios remain fairly constant whether the total number of crosses is 50, 100, 200, or 400 (Fig. 2A). Changes in total experiment size shift all the curves proportionately. Thus, with balanced, nested designs the number of crosses should be maximized and as many degrees of freedom as possible should be allocated to the male or half-sib family mean square by using two females per male. Using only one female per male would not allow estimation of $\sigma_A^2$ separate from $\sigma_D^2$. Only if non-additive genetic variances are large should much effort be made to increase the numbers of females (full-sibs) per male (half-sib), and then not much beyond six.

In contrast, the factorial design is generally satisfactory for estimating $\sigma_A^2$ for any allocation of 100 crosses into half- or full-sib groups (Fig. 2B). Equal numbers of male and female parents are most efficient. This allocation creates the maximum number of full-sibs per half-sib family. This is an exception to the general experience that greater efficiencies for estimating $\sigma_A^2$ occur with maximum numbers of half-sibs. However, in terms of allocating degrees of freedom, all allocations of the 100 crosses examined provided a minimum of 18 d.f. for the mean square estimating the half-sib variances. This value is greater than that for some of the poorer allocations in the nested design.

The more crosses made in the factorial design, the smaller the $V(\hat{\sigma}_f^2)/\sigma_A^2$ ratio (Fig. 3B) but the differences are not as marked as they were for the nested design. With at least six parents of either sex and up to equal allocations of males and females, even as few as 50 crosses in an experiment can be satisfactory.

Even greater efficiencies are afforded by the partial diallel designs which can create more d.f. for the half-sib variances (general combining ability) than the other balanced designs for the same number of crosses (Fig. 2C). As is the case for the other designs, larger experiment sizes create greater efficiencies but the scale of improvement is more marked than for the factorial design (Fig. 3C).

If $\sigma_D = 0$, the number of crosses is fixed at 100, and only efficient male-female allocations are used, the nested design (A/B) is poorest, the factorial (AB) next, and the diallel (AA) is most efficient for estimating $\sigma_A^2$ (Fig. 4).

Estimating $\sigma_D^2$

When dominance genetic variance is present and is to be estimated, the relative ranking of the allocations and designs will change from that of estimating $\sigma_A^2$ because the mean square estimators and their degrees of freedom have different effects. For the following comparisons we assume $\sigma_D^2 = \sigma_A^2$. The efficiency curves for the nested design (A/B) do not markedly differ for estimating $\sigma_D^2$ than for estimating $\sigma_A^2$ except when the number of females per male drops below four and most especially at low levels of $\sigma_A^2$ and $\sigma_D^2$ (Fig. 5A). Three or fewer females per male decreases the efficiency of estimating the full-sib variances (MSF/M) more than offsetting the efficiency for estimating the half-sib variances (MSM) which is used to estimate $\sigma_D^2$.

Since the estimation of $\sigma_D^2$ for the factorial and diallel designs requires only the error and full-sib mean squares,
most efficient (Fig. 6). However, the diallels with s values around six (six crosses per half-sib family) are nearly as efficient.

Since high levels of $\sigma_1^2$ would also change the mean squares used to estimate $\sigma_\lambda^2$, all of the efficiencies for estimating $\sigma_1^2$ decline (Fig. 7) as $\sigma_1^2$ increases from 0 to $\sigma_\lambda^2$. The relative rankings remain generally the same but at low levels of $\sigma_1^2$ and $\sigma_\lambda^2$ the diallels with s values greater than 3 are more consistently better than the factorials.

**Fixing Numbers of Parents**

Slightly different comparisons can be drawn if the number of parents is fixed. This might occur if external conditions such as effective population size limitations are imposed for breeding purposes and the parental sample is fixed. Under these conditions, the nested design affords little flexibility. With 100 parents, the best that can be done is to use 33 males with 2 females per male (99 parents in 66 crosses) omitting one line for balance. The factorial and diallel design allocations permit many more mating patterns and numbers of crosses. Assuming 300 to 600 crosses for the 100 parents, factorial designs generally have lower errors of estimation than nested designs but higher errors than diallels (Fig. 8).

Doubling the number of crosses

their efficiencies are mainly dependent on the degrees of freedom available for the full-sib (MSFM and MSS) mean squares. Thus, for the factorial (AB) the same ranking for $\delta_\lambda^2$ as for $\delta_1^2$ exists since both allocate most degrees of freedom to MSFM when equal numbers of males and females are used (Fig. 5B). For the diallel (AA) the exact reversal of efficiency ranking for estimating $\delta_1^2$ than for estimating $\delta_\lambda^2$ exists (Fig. 5C). With 100 crosses, when $\sigma_1^2 = \sigma_\lambda^2$ the factorial design, which allocates so many degrees of freedom to the full-sib (MSMF) mean square, is reduces error more in the factorial design than in the diallel (Fig. 8).

**Blocking**

In large experiments replicates often become unreasonably large and sub-blocking is often desirable both for estimating means and variances. Sub-blocking of replicates can be imposed on the different designs with varying degrees of ease. The easiest is the blocked disconnected diallel, but all designs can be sub-blocked. The appropriate
Figure 6. — Relative design efficiencies for estimating $\sigma_p^2 (V(\delta_p^2)/\sigma_p^2)$ when $\sigma_p^2 = \sigma_A^2$, $c = 100$.

Figure 7. — Relative design efficiencies for estimating $\sigma_A^2 (V(\delta_A^2)/\sigma_A^2)$ when $\sigma_A^2 = \sigma_A^2$, $c = 100$.

Figure 8. — Design comparisons of efficiency $(V(\delta_A^2)/\sigma_A^2)$ when number of parents is constant (100), for $\sigma_p^2 = 0$.

methods are by separating male groups in the nested design, by separating different male and female groups without double entry in the factorial, by separating "chains" of crosses in the partial diallel, and by disconnecting blocks in the disconnected diallels. In addition, some overlapping of genetic entries can be allowed between books, but this alternative is ignored here.

When soil is heterogeneous as expressed by a decline from one, H. F. Senn's (1938) coefficient (b), large replications engender increasing plot errors. Consequently, errors in estimating means and genetic variances also increase. By sub-blocking, the effects of large plot errors can be mitigated, but at some cost in the degrees of freedom for estimating genetic variances. In Fig. 9 we compare the effects of creating four sub-blocks out of 100 crosses for some permissible design allocations. When $b = 0.5$, the $V(\delta_A^2)/\sigma_A^2$ curve for unit $\sigma_p^2$ (but larger $\sigma_A^2$) is higher than when $b = 1$ and when blocking is not used. When four sub-blocks are created however, the curve generally drops but remains higher than the curve for homogeneous soil conditions. At lower levels of $b$ (greater heterogeneity), the advantage of blocking would be greater. At higher levels of $b$ the loss of degrees of freedom caused by sub-blocking exceeds the gain from reducing the plot error variance.

Figure 9. — Comparison of effects of sub-blocking for four design allocations with $c = 100$, $b = .5$, $\sigma_p^2 = 0$, and $p = 1$ and 4.
In the diallels examined, the value of blocking is never very great. In the nested design, when heterogeneity exists (b = 0.5), and no sub-blocking is used, the variance curve for 16 males and six females per male is pushed closer to the CV = ½ curve at high heritabilities and is higher than that standard at heritabilities less than 0.25. However, blocking is futile because of the loss of necessary degrees of freedom for males. Blocking is most inefficient in the factorial design with eight and 12 parents of alternate sex. While the unblocked design is not much worse when b = 0.5 than when b = 1, blocking reduces the critical degrees of freedom so much that it is less efficient than the nested design at heritabilities below 0.4.

**Conclusion**

Choice of type of balanced mating design, allocation of effort between numbers of half-sibs versus full-sibs, increasing total experiment size, and blocking have all been examined under certain conditions over wide ranges of heritability and dominance variance. The equations developed in this paper permit direct examination of efficiency in estimating genetic variances when those factors are varied. Diallel and factorial designs usually are more efficient for estimating both dominance and additive genetic variances. In nested and diallel designs, most effort should be placed on creating large numbers of half-sib families, and no more than six full-sib families per half-sib would be useful. In the factorial design, equal numbers of males and females are almost always desirable. For all cases, increasing the total experiment size is very effective in increasing estimation efficiency. However, the factorial design uses equal numbers of males and females best and is less affected by increases in total experiment size than the other designs. The nested design requires at least 100 crosses, well allocated among half- and full-sibs. For the factorial and diallel designs, 50 crosses may be sufficient for a wide range of allocations. The advantages of blocking are surprisingly ambiguous. Blocking is not generally beneficial for variance estimation though it might well be useful for other purposes.

**Summary**

The efficiency of four mating designs for estimating genetic variances is examined over a continuous range of heritabilities and for varying conditions of allocations, experiment size, levels of dominance genetic variance, and blocking. Explicit equations are developed which allow for very general examination of these balanced mating designs. 

*Key words: Genetic variance, variance components, heritability, analysis of variance, mating design.*

**Literature Cited**


**Appendix**

Appendix Table A-1

<table>
<thead>
<tr>
<th>Sampling Variance for Estimates of ( \sigma_{a}^{2} ) and ( \sigma_{d}^{2} ) in Terms of Mean Squares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocked Nested (AA/AA)</td>
</tr>
<tr>
<td>( \sigma_{a}^{2} ) = ( \frac{32}{r^{2}} \left( \frac{3}{1} \right) (\sum_{i} \bar{x}<em>{i}^{2}) - (\sum</em>{i} \bar{x}_{i})^{2} )</td>
</tr>
<tr>
<td>Blocked Factorial (AB)</td>
</tr>
<tr>
<td>( \sigma_{a}^{2} ) = ( \frac{r^{2}}{r^{2}} \left( \frac{32}{r^{2}} \right) \left( \sum_{i} \bar{x}<em>{i}^{2} \right) - (\sum</em>{i} \bar{x}_{i})^{2} )</td>
</tr>
<tr>
<td>Blocked Partially Balanced (AA - Partial)</td>
</tr>
<tr>
<td>( \sigma_{a}^{2} ) = ( \frac{r^{2}}{r^{2}} \left( \frac{32}{r^{2}} \right) \left( \sum_{i} \bar{x}<em>{i}^{2} \right) - (\sum</em>{i} \bar{x}_{i})^{2} )</td>
</tr>
<tr>
<td>Blocked Disconnected Partially Balanced (AA - Disconnected)</td>
</tr>
<tr>
<td>( \sigma_{a}^{2} ) = ( \frac{r^{2}}{r^{2}} \left( \frac{32}{r^{2}} \right) \left( \sum_{i} \bar{x}<em>{i}^{2} \right) - (\sum</em>{i} \bar{x}_{i})^{2} )</td>
</tr>
</tbody>
</table>

50
Appendix Table A.2

Sampling Variances for Estimates of $\sigma^2_a$ and $\sigma^2_d$ in Terms of Heritability and the Dominance Ratio

**Blocked Nested (A/B)**

\[
\frac{v(a^2)}{v} = (2\overline{g}^2/2(n-p)r_1 r_2) + (2\overline{g}^2/2n) r_3 r_4 + \frac{1}{n} \overline{y}_4 [r_7 + \overline{g}_6] + \frac{1}{1-n} 2\overline{y}_5 [r_7 + \overline{g}_6] + \frac{1}{(1-n)(1-\gamma)} 2\overline{y}_5 [r_7 + \overline{g}_6] + \frac{1}{(1-n)(1-\gamma)} \overline{y}_5 r_2 + \frac{1}{(1-n)(1-\gamma)} \overline{y}_5 r_5 + \frac{1}{1-n} \overline{y}_6 + \frac{1}{n} \overline{y}_6
\]

\[
\gamma_1 = \text{ref}^2, \quad \gamma_2 = \text{mf}(\overline{r}-1), \quad \gamma_3 = \text{mf}-p, \quad \gamma_4 = \frac{3\text{ref}^2(1-\overline{r})}{\text{mf}^2} + 4\overline{r}^2(\text{ref}^2-1) - \text{mf}^2(\text{ref}^2-1) + \text{mf}^2 \overline{r}\text{ref}^2
\]

\[
\gamma_5 = 2n-1, \quad \gamma_6 = \text{ref}^2
\]

**Appendix Table A.2 (Continued)**

\[
\frac{v(d^2)}{v} = (2\overline{g}^2/2n) + (8\overline{g}^2\overline{g}_5 [r_{11} + \overline{g}_9] + \frac{1}{1-n} 8\overline{g}^2 \overline{g}_7 [r_{11} + \overline{g}_9] + \frac{1}{n} \overline{y}_4 [r_7 + \overline{g}_6] + \frac{1}{1-n} 2\overline{y}_5 [r_7 + \overline{g}_6] + \frac{1}{(1-n)(1-\gamma)} 2\overline{y}_5 [r_7 + \overline{g}_6] + \frac{1}{(1-n)(1-\gamma)} \overline{y}_5 r_2 + \frac{1}{(1-n)(1-\gamma)} \overline{y}_5 r_5 + \frac{1}{1-n} \overline{y}_6 + \frac{1}{n} \overline{y}_6
\]

\[
\gamma_1 = \text{ref}^2, \quad \gamma_2 = \text{mf}(\overline{r}-1), \quad \gamma_3 = \text{mf}-p, \quad \gamma_4 = \frac{3\text{ref}^2(1-\overline{r})}{\text{mf}^2} + 4\overline{r}^2(\text{ref}^2-1) - \text{mf}^2(\text{ref}^2-1) + \text{mf}^2 \overline{r}\text{ref}^2
\]

\[
\gamma_5 = 2n-1, \quad \gamma_6 = \text{ref}^2
\]
Appendix Table A.2 (Continued)

**Blocked Factorial (AB)**

\[
V_{ab}^{2} = \frac{(2 \cdot p^{2} \cdot q^{2} \cdot r^{2})}{(2a_{1} \cdot 2a_{2} \cdot 2r)} + \frac{1}{2} \cdot p^{2} \cdot q^{2} \cdot r \cdot s + \frac{1}{2} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \cdot 2r \cdot 2s
\]

\[
\begin{align*}
&+ \frac{1}{(1-1)} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \\
&+ \frac{1}{(1-1)} \cdot \frac{1}{(1-1)} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \\
&+ \frac{1}{(1-1)} \cdot \frac{1}{(1-1)} \cdot \frac{1}{(1-1)} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \cdot 2r
\end{align*}
\]

\[
a_1 = f(\text{AB}) \cdot (r-p) + m(\text{AB}) \cdot (r-p),
\]

\[
a_2 = 4a_2^2 \cdot (2a_1^2 - 1) \cdot (2r - 1) \cdot m^2,
\]

\[
a_3 = a_1 \cdot a_2 \cdot (r-p) \cdot (r-p)^2 + p \cdot (r-p) \cdot 2r \cdot 2s,
\]

\[
a_4 = r \cdot m, a_5 = (2a_1^2 - 1) \cdot (2r - 1) \cdot m^2,
\]

\[
a_6 = a_1 + a_2 + a_3 + a_4 + a_5 + a_6^2 \cdot (r-p) \cdot (r-p)^2 + a_7^2 \cdot (r-p) \cdot (r-p)^2,
\]

\[
a_7 = f(\text{AB}) + m(\text{AB}) \cdot (r-p) \cdot (r-p)
\]

\[
V_{ab}^{2} = 32 \cdot p^2 \cdot q^2 \cdot r^2 \cdot (\frac{1}{2} \cdot p^2 \cdot q^2 \cdot r^2 + \frac{1}{2} \cdot p^2 \cdot q^2 \cdot r^2 + \frac{1}{2} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \cdot 2r \cdot 2s)
\]

\[
\begin{align*}
&+ \frac{1}{(1-1)} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \\
&+ \frac{1}{(1-1)} \cdot \frac{1}{(1-1)} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \\
&+ \frac{1}{(1-1)} \cdot \frac{1}{(1-1)} \cdot \frac{1}{(1-1)} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \cdot 2r
\end{align*}
\]

\[
f_1 = 4a_1, f_2 = 2a_2, f_3 = 2a_3, f_4 = 2a_4,
\]

\[
\sum_{i=1}^{n} f_i = n^2 + 1, f_5 = a_5 \cdot (r-p), f_6 = a_6 \cdot (r-p)
\]

Appendix Table A.2 (Continued)

**Blocked Partial Diallel (AA - Partial)**

\[
V_{aa}^{2} = \frac{1}{2} \cdot p^2 \cdot q^2 \cdot r^2 \cdot (\frac{1}{2} \cdot p^2 \cdot q^2 \cdot r^2 + \frac{1}{2} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \cdot 2r \cdot 2s + \frac{1}{(1-1)} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \cdot 2r \cdot 2s + \frac{1}{2} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \cdot 2r \cdot 2s)
\]

\[
\begin{align*}
&+ \frac{1}{(1-1)} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \\
&+ \frac{1}{(1-1)} \cdot \frac{1}{(1-1)} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \\
&+ \frac{1}{(1-1)} \cdot \frac{1}{(1-1)} \cdot \frac{1}{(1-1)} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \cdot 2r
\end{align*}
\]

\[
c_1 = 4a_1^2 \cdot (2a_1^2 - 1) \cdot (2r - 1) \cdot m^2, c_2 = 2a_1, c_3 = m^2 + 2a_1, c_4 = q \cdot (2a_1 - 2r - 1),
\]

\[
c_5 = q \cdot (2a_1 - 2r - 1), c_6 = q \cdot a_1 \cdot (2a_1 - 2r - 1)
\]

\[
V_{aa}^{2} = (64 \cdot p^2 \cdot q^2 \cdot r^2) \cdot (\frac{1}{2} \cdot p^2 \cdot q^2 \cdot r^2 + \frac{1}{2} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \cdot 2r \cdot 2s)
\]

\[
\begin{align*}
&+ \frac{1}{(1-1)} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \\
&+ \frac{1}{(1-1)} \cdot \frac{1}{(1-1)} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \\
&+ \frac{1}{(1-1)} \cdot \frac{1}{(1-1)} \cdot \frac{1}{(1-1)} \cdot a_{1} \cdot a_{2} \cdot 2p \cdot 2q \cdot 2r
\end{align*}
\]

\[
f_1 = 2a_1, f_2 = 2a_2, f_3 = 2a_3, f_4 = 2a_4,
\]

\[
f_5 = f(\text{A}) + m(\text{A}) \cdot (r-p) \cdot (r-p), f_6 = a_6 \cdot (r-p)
\]

52
Beiträge zum geographischen Variationsmuster der Douglasie

Von K. Stern

Zusammengestellt von A. König1) und H. H. Hattemer2)

(Eingegangen im Februar 1974)

Einleitung

Im Rahmen des Emslandprogrammes, eines Projektes zur Verbesserung der wirtschaftlichen Möglichkeiten in einem Landstrich in Nordwest-Deutschland, wurden im Forstamt Nordhorn im Bereich der Landwirtschaftskammer Weser-Ems vom Institut für Forstgenetik und Forstpflanzenzüchtigung drei Herkunftsversuche mit Douglasien angelegt. Das Ziel war, die Erfahrungen über den Erfolg von Anbauten mit verschiedenen Provenienzen dieser Baumart zu ergründen.

1) Institut für Forstgenetik und Forstpflanzenzüchtigung der Bundesforstwirtschaft, 2240 Großhansdorf 2 (Schmalenbeck), Sickerlandstraße 2.
2) Abteilung für Forstliche Biometrie der Universität, 3400 Göttingen, Bürgenweg 5.

Die Anlage, sowie die erste Auswertung der Feldversuche wurden unter Anleitung von K. Stern durchgeführt. Die im folgenden mitgeteilten Ergebnisse basieren auf seinem Bericht an die Deutsche Forschungsgemeinschaft.

Material und Methoden

Das Saatgut wurde von forstgenetischen Institutionen der USA und Kanadas bezogen, die es selbst in autochthonen Beständen für eigene Versuche eingesammelt hatten. Die Samenproben stammen aus dem gesamten Verbreitungsgebiet der Douglasie, die Herkunftgebiete sind jedoch unterschiedlich stark vertreten. Auch ist die Zahl der Bäume, von denen Zapfen geerntet wurden, pro Herkunft oft sehr gering.

Silvae Genetica 23, 1-3 (1974)