Statistical Efficiency of Row and Noncontiguous Family Plots in Genetic Tests of Loblolly Pine

BY C. C. LAMBETH1), W. T. GLADSTONE2) and R. W. STONECIPHER3)
Weyerhaeuser Company
Forestry and Timber Products Research Division

(Received 12th March 1982)

Summary

Eight genetic tests comparing five-tree family row plots and five-tree noncontiguous plots (trees within a plot are located at random throughout the replicate-block) were analyzed for fourth- or fifth-year tree height. Compared to the row arrangement, the noncontiguous plots resulted in much lower coefficients of variation for family means since block-by-family interactions were essentially eliminated. Row plots had large block-by-family interactions and higher coefficients of variation. No effects of plot arrangement on within-plot variance were detected. Twelve blocks of noncontiguous plots would estimate a family mean as precisely as 17 blocks of row plots.

Despite the value of smaller tests or increased precision, especially on sites with extreme environmental variability, noncontiguous plots are somewhat more difficult to install than row plots and the tracking of individual trees throughout measurement cycles demands concentration. There may also be concern about the effects of inter-versus intra-family competition on family and individual ranking and variance estimates.

Key words: Genetic Testing, Row Plots, Noncontiguous Plots, Competition, Coefficient of Variation, Pinus taeda.

Zusammenfassung


Es werden die Vor- und Nachteile sowohl der nicht-zusammenhängenden als auch der Reihenparzellen diskutiert. Bei jenen kann man insbesondere auf uneinheitlichen Standorten eine höhere Genauigkeit erzielen und dadurch den Versuchsumfang (Anzahl der Bäume) verringern. Als Nachteil wird hingegen der höhere Aufwand bei der Identifizierung der einzelnen Bäume sowohl bei der Versuchsplanung als auch bei der späteren Erhebungen angesehen. Außerdem können unterschiedliche Konkurrenzeffekte zwischen und innerhalb von Familien bestehen, wodurch die Ränge von Familien und Individuen und die Schätzungen für die Varianzen beeinflusst werden könnten.

Introduction

Genetic tests are one of the most costly phases of tree improvement. They should comprise efficient designs which are no larger than is necessary to achieve the test objectives and to cover anticipated mortality. A test that is too large will take in extraneous site variability and will involve the costs of carrying unnecessary individuals from controlled pollination through analysis of the last field measurements.

Much of the discussion regarding genetic test design has revolved around plot shape and size with the assumption that a Randomized Complete Blocks design would be employed and families are the only "treatments" (Conkle 1963, Evans et al. 1961, Shure and Faust 1961, Stonecipher 1966, Wright 1976). The general conclusion is that small plots are most efficient when efficiency is expressed in terms of the number of individuals needed to achieve a defined accuracy in the estimate of a family mean.

The purposes of blocking are to:
1) ensure that members of the family sample the environmental variation within the test area and thereby reduce environmental error in estimating the family mean, i.e., the mean is a better reflection of the genetic worth of the family and
2) account for differences among blocks in the analysis of variance.

The single-tree plot is best able to achieve these purposes (Conkle 1963, Wright 1976) because the number of microsites sampled with a given number of trees per family is maximized and, since block size is minimized variance within blocks is minimized while that between blocks is maximized.

Single-tree plots are rarely used in genetic testing in spite of their statistical efficiency for at least five commonly stated reasons:

1. Uncertainty may exist over inter- versus intra-family competition effects on family and individual-within-family rankings and variance among families. Single-tree plots maximize intrafamily competition and minimize intrafamily competition while the reverse is true for square plots with buffer rows. This disadvantage also applies to row plots, though to a lesser extent.
2. Within family selection may be hampered since sibs are not in the more localized environment encountered in a row or square plot. This problem is critical only if within plot comparison of sibs is an important part of the within-family selection process.
3. Thinning by family in genetic tests would be more difficult. For example, thinning each family to a certain number of trees and at the same time achieving
reasonably uniform spacing over the entire test would be simpler with multiple-tree plots.

4. They are somewhat more difficult to lay out and track over time. Since they are more prone to misidentification, each tree should be tagged.

5. Analyses may be more difficult if computing software is not available to deal with inevitable missing plots. Furthermore, a plot mean analysis of variance is often used to reduce computing costs in tests of multiple-tree plots.

The noncontiguous arrangement is a multi-tree alternative to the single-tree plot; wherein family plot members are located at random within the block rather than contiguously in rows or squares. It has the advantage of the single-tree plot in that microsite variation in the test area is thoroughly sampled while the inclusion of several trees alleviates the threat of missing plots. The intent of the noncontiguous arrangement is to ensure that plot members are exposed to the environmental variability within the block just as blocking is done to ensure that plots of the family are exposed to all of the environmental variability within the test area.

Noncontiguous plots have had little exposure or use in genetic testing in forestry although they have recently been advocated by Libby and Cockshram (1981). Most researchers are using small row plots (10 or fewer trees). In this report we compare analytical results of row and noncontiguous plots of equal size.

Materials and Methods

Eight genetic tests (Table 1) having side-by-side comparisons of blocks with family row plots and blocks with noncontiguous plots were analyzed for tree height at age four or five years. The design was Randomized Complete Blocks, usually with eight blocks containing family row plots and four with noncontiguous plots. For the plot type comparison, each test was divided into three sets of four replications — two with row plots and one with noncontiguous plots. The sets were constructed to ensure that a set of blocks with noncontiguous plots would be of equal area to either set of blocks with row plots with which it could be compared. Two tests had only eight blocks a set of four blocks each with row and noncontiguous plots.

An analysis of variance (Table 2) was conducted for all sets at each test. The degrees of freedom were identical for sets of a given test except for the within-plot degrees of freedom which varied slightly due to differential mortality. Five trees were planted in all plots. Variance components were obtained from the expected mean squares.

Table 2. — Form of the analysis of variance conducted for each set in each test described in Table 1.

<table>
<thead>
<tr>
<th>Source</th>
<th>d.f.</th>
<th>Mean Square</th>
<th>t-test</th>
<th>Expected Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocks</td>
<td>(b-1)</td>
<td>MS_b</td>
<td>MS_b/MSE</td>
<td></td>
</tr>
<tr>
<td>Families</td>
<td>(f-1)</td>
<td>MS_f</td>
<td>MS_f/MS_BF</td>
<td></td>
</tr>
<tr>
<td>Block x Family</td>
<td>(b-1)</td>
<td>MS_BF</td>
<td>MS_BF/MS_f</td>
<td></td>
</tr>
<tr>
<td>Within Plots</td>
<td>(b-1)</td>
<td>MS_y</td>
<td>MS_y/MS_BF</td>
<td></td>
</tr>
</tbody>
</table>

\[ y^2 = n \cdot s_t^2 + n \cdot s_{BF}^2 + s_y^2 \]

\[ s_{BF}^2 = s_{BF}^2/n \]

Where: \( n \) = number of blocks, \( f \) = number of families, \( s_t^2 \) = mean square for the ith family in the jth block

Note: \( s_y^2 \) will vary in the Expected Mean Squares depending on sample variation by plot, block, and family.

Table 2(b) — Variance components for tree height in 4- and 5-year-old tests from data sets with row and noncontiguous family plots in the same test. There were four blocks per data set and five trees planted per family plot in each block.

\[ s_{BF}^2 = s_{BF}^2/n \]

\[ \frac{1}{n} y^2 = \text{Family variance} \]

\[ s_{BF}^2 = \text{Block-by-family interaction variance} \]

\[ * \text{Significant at the .05 probability level.} \]

A plot mean coefficient of variation was calculated for each set as follows:

\[ \text{C.V. (Plot Means)} = \sqrt{\frac{\text{MS}_\text{BF}}{\text{MS}_\text{B} \cdot n_H}} \]

Where: \( n_H \) = Harmonic mean number of surviving trees per plot

\( X \) = Overall mean of the set

Table 1. — Description of the genetic tests used in comparisons of row and noncontiguous plots.

<table>
<thead>
<tr>
<th>Genetic Test</th>
<th>Planting Location</th>
<th>Number of Families</th>
<th>Family Type</th>
<th>Number of Blocks with Row Plots</th>
<th>Number of Blocks with Noncontiguous Plots</th>
<th>Age-Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Newfort Co., N.C.</td>
<td>24</td>
<td>FS</td>
<td>4 (1 set)</td>
<td>4 (1 set)</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>Greene Co., Ala.</td>
<td>24</td>
<td>FS</td>
<td>8 (2 sets)</td>
<td>4 (1 set)</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>Kenerly Co., Miss.</td>
<td>24</td>
<td>FS</td>
<td>8 (2 sets)</td>
<td>4 (1 set)</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>Greene Co., Ala.</td>
<td>24</td>
<td>OP</td>
<td>8 (2 sets)</td>
<td>4 (1 set)</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>Greene Co., Ala.</td>
<td>24</td>
<td>FS</td>
<td>8 (2 sets)</td>
<td>4 (1 set)</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>Garland Co., Al</td>
<td>20</td>
<td>OP</td>
<td>8 (2 sets)</td>
<td>4 (1 set)</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>Savier Co., Al</td>
<td>18</td>
<td>OP</td>
<td>8 (2 sets)</td>
<td>4 (1 set)</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>Howard Co., Al</td>
<td>17</td>
<td>OP</td>
<td>8 (2 sets)</td>
<td>4 (1 set)</td>
<td>4</td>
</tr>
</tbody>
</table>

*FS = Full-sib, OP = Open-pollinated
Table 4. — Plot mean coefficient of variation (C.V.) and the corresponding number of blocks (b) needed to estimate a family mean height within an allowable error of ± 3% with a 9% risk that the allowable error would be exceeded. The results are from four- and five-year-old genetic tests with five trees per family plot and four blocks per data set with side-by-side comparisons of row and noncontiguous plot designs.

<table>
<thead>
<tr>
<th>Test-Set</th>
<th>Row Plots</th>
<th></th>
<th>Noncontiguous Plots</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C.V.</td>
<td>b</td>
<td>C.V.</td>
<td>b</td>
</tr>
<tr>
<td>1-1</td>
<td>.065</td>
<td>9</td>
<td>.083</td>
<td>13</td>
</tr>
<tr>
<td>2-1</td>
<td>.109</td>
<td>20</td>
<td>.092</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>.108</td>
<td>20</td>
<td>.091</td>
<td>15</td>
</tr>
<tr>
<td>3-1</td>
<td>.112</td>
<td>21</td>
<td>.084</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>.109</td>
<td>20</td>
<td>.095</td>
<td>16</td>
</tr>
<tr>
<td>4-1</td>
<td>.102</td>
<td>18</td>
<td>.091</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>.084</td>
<td>13</td>
<td>.084</td>
<td>13</td>
</tr>
<tr>
<td>5-1</td>
<td>.123</td>
<td>25</td>
<td>.088</td>
<td>12</td>
</tr>
<tr>
<td>6-1</td>
<td>.065</td>
<td>9</td>
<td>.058</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>.061</td>
<td>8</td>
<td>.063</td>
<td>8</td>
</tr>
<tr>
<td>7-1</td>
<td>.098</td>
<td>17</td>
<td>.063</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>.094</td>
<td>16</td>
<td>.082</td>
<td>13</td>
</tr>
<tr>
<td>8-1</td>
<td>.089</td>
<td>14</td>
<td>.082</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>.136</td>
<td>30</td>
<td>.082</td>
<td>13</td>
</tr>
</tbody>
</table>

Average of Tests .096 17 .081 12

The plot mean C.V. was used to determine the number of blocks of five-tree row plots needed to estimate a family mean within an allowable ± error and at a defined probability level. If the allowable error is expressed as a fraction of the mean then the appropriate equation, adapted from Snedecor and Cochran (1967), is:

\[ b = \frac{t(C.V. Plot Means)^2}{d^2} \]

Where:  
- \( b \) = number of blocks needed  
- \( t \) = t-table value for the appropriate confidence level.  
  (We accepted a 5% risk that the allowable error would be exceeded.)  
- \( d \) = allowable error in the estimate of the mean.  
  (We accepted a 5% error error in the estimate of a family mean height.)

The values (b and t) are not independent. The tabular value of (t) needed for the equation depends on the degrees of freedom associated with (b). The final solution for (b) was obtained interactively from the t-table and intermediate solutions of (b). The test C.V. is, in a sense, the average C.V. for all families in the test. In using this C.V. to estimate (b) it was necessary to assume that families within a test would have the average C.V.

Results and Discussion

Block-by-family interaction could be expected to be more serious in row plots than in noncontiguous plots. Trees within a row plot often experience a common environment effect which causes the mean to deviate from the actual genetic value obtained from a family planted throughout the block. Thus, a family's plot mean relative to other families can vary dramatically from block to block.

The results indicate that there was greater block-by-family interaction in row plots than in noncontiguous plots (Table 3); no statistically significant interaction of that type was detected with noncontiguous plots. The block-by-family variance component was near zero in all noncontiguous data sets whereas the same component for row plots was often significant and as large or larger than the family variance component from the same data set.

Since variance components are usually correlated with total variance in the test area, plotting the family, block-by-family and within plot variance components against the total variance of each data set should graphically demonstrate major differences between row and noncontiguous plots. The graphs reveal that the noncontiguous arrangement affected only the block-by-family interaction in a dramatic way (Figures 1—3). Differences between row and noncontiguous plots for family and within plot variance components were not demonstrable. An increase in within plot variance with noncontiguous rather than row plots would have been expected if a common environment effect was the cause of greater block-by-family interaction in the row plots. The mean of the within-plot variances for the noncontiguous plots was slightly larger though total variances were slightly larger in the row plot configuration. This suggests that there was a small increase in within plot variance with noncontiguous plots as could logically be expected.

Figure 1. — Family variance component plotted against total variance for several data sets with row or noncontiguous plots. The data were from four- and five-year-old genetic tests of loblolly pine.
The coefficient of variation (equation 1) provides a convenient method to compare the overall effect of the two plot types on test precision since the MS_{BXF} Expected Mean Squares includes all within family variance, both within plot and block-by-family interaction variances. Noncontiguous plot coefficients of variation were almost always smaller than those for row plots.

Trees in three of the tests were large enough to yield volume data (tests 6, 7 and 8). Analyses of these data gave almost identical results to those for tree height in that C.V.’s were low and block-by-family interactions were nonsignificant in noncontiguous plots.

The number of blocks (b) needed to estimate a family mean with defined precision is a function of the square of the C.V.. Consequently, the noncontiguous plot appears to be much more precise, requiring only 12 blocks versus 17 for the row-plot design. For a test of 35 families with two border rows and five trees per plot, the row plot design would require about 2975 test trees and 453 border trees while the noncontiguous design would require only 2100 test trees and 383 border trees. At a 3 m × 3 m spacing the respective land areas required would be 3.1 and 2.2 hectares.

Obvious advantages of noncontiguous plots over row plots are related to test size differences. With a smaller test:

1. Fewer controlled matings are required to produce the seed.
2. Fewer plants need to be grown for outplanting.
3. Uniform sites for field testing would be easier to find since the required land area is smaller.
4. Fewer trees would be outplanted, measured and maintained.
5. Fewer data to be processed.
6. Nonplantable spots in the test area can be dealt with more effectively.

**Conclusions**

Noncontiguous plots were effective in eliminating block-by-family interaction in the eight tests examined whereas the same term was highly significant in row plots. Consequently, coefficients of variation were lower and the number of trees required to achieve a defined precision in the estimate of a family mean was considerably lower for non-contiguous than for row plots — 60 versus 85 in an average...
four our five-year-old lobolly pine genetic test with five-tree plots. Within-plot variances for noncontiguous plots may have been slightly larger than those for row plots though it was no possible to statistically confirm this conclusion.

The significant advantages that the commonly used row plot may have over the noncontiguous arrangement are ease of layout and tracking in the field, and simplicity in silvicultural thinning by family. The efficiency advantage of noncontiguous plots, i.e., fewer trees needed to realize the same precision in a genetic test, seems to outweigh the advantages of row plots.

Acknowledgements

We wish to thank Claude O'Gwynn and Dick Welch who were instrumental in the establishment, maintenance and measurement of the tests, Barbara Bower who constructed the figures, David Pohl who assisted in computer analyses and Cherri Walton who typed the manuscript.

Literature Cited


Methods of estimating the average performance of families across incomplete open-pollinated progeny tests

By P. P. Cotterill1), R. L. Correll2) and R. Boardman3)

(Received 15th March 1982)

Summary

Six mathematical procedures are outlined and compared for relative accuracy in estimating the average performance of families from open-pollinated progeny data which are incomplete in the sense that families are represented in some but rarely all progeny tests. In this instance the data were records of stem volume and stem straightness (five-point visual score) from five open-pollinated progeny tests of Pinus radiata in South Australia. The methods were rank-score (RS), site-adjustment (SA) or standard site-adjustment (SSA) procedures compared with least-squares (LS), weighted least-squares (WLS) or shrunken least-squares (SLS) procedures. Logarithmic transformation was used to stabilise the variance of volume across sites.

The LS, WLS and SLS methods agreed very closely in their evaluation of families for all traits studied. The RS and SSA evaluations of families for volume agreed reasonably closely with least-squares evaluations, while the SA evaluations for volume were less accurate. All methods provided essentially the same evaluation of families for straightness.

Key words: progeny testing, non-orthogonality, genotype value, log transformation, least-squares.

Zusammenfassung


Dabei wurde festgestellt, daß die LS-, WLS- und SLS-Verfahren sehr weit beieinander der Wachstumsleistungen aller untersuchten Familien übereinstimmten. Die RS- und SSA-Berechnungen für die Stammvolumen der jeweiligen Familien stimmten hinreichend mit der Auswertung der kleinsten Quadrate überein, während die SA-Ermittlungen der Volumen weniger genau waren. Dagegen ergaben sämtliche Berechnungsarten im wesentlichen dieselben Er- gebnisse für die Ermittlung der Stammgeradenheit der Fa- milien.

Introduction

Most tree breeding programs rely on progeny testing in each generation to determine the genetic merit of new selections. These tests are usually established over a number of sites, and the analysis of data is primarily directed towards estimating the average relative performance of families across the range of conditions. When the family × progeny test array of available data is complete (all families represented at all sites) the ana- lysis is straightforward in that family performance can be obtained by simply averaging across sites. More often, the family × progeny test table is incomplete with a propor- tion of families represented at only a few sites, and in this case the non-orthogonality of the data can complicate