

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

By C. C. LAMBETH<sup>1)</sup>, W. T. GLADSTONE<sup>2)</sup>  
and R. W. STONECYPHER<sup>3)</sup>

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

In acht genetischen Tests (Feldversuchen) mit *Pinus taeda* wurde die Effizienz von 5-Baum-Reihenparzellen gegenüber nicht-zusammenhängenden 5-Baum-Parzellen untersucht. In letzteren stehen die Bäume nicht benachbart, sondern sind zufallsmäßig über die Wiederholung (Block) verteilt. Als Daten wurden die Höhen im Alter von vier bzw. fünf Jahren verwendet. Die nicht-zusammenhängenden Parzellen ergaben im Vergleich zur Reihenanordnung viel niedrigere Variationskoeffizienten für Familienmittelwerte, da die Wechselwirkungen für Blöcke  $\times$  Familien eliminiert wurden. Reihenparzellen hatten hohe Interaktionen für Blöcke  $\times$  Familien und höhere Variationskoeffizienten. Hingegen wurde keine durch den Parzellentyp bedingte Auswirkung auf die Varianz innerhalb von Parzellen festgestellt. Ein Familienmittel würde bei nicht-zusammenhängenden Parzellen mit 12 Blöcken ebenso genau geschätzt, wie bei 17 Blöcken mit Reihenparzellen.

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 Versuchsbegründung als auch bei den späteren Erhebungen angesehen. Außerdem können unterschiedliche Konkur-

renzeffekte zwischen und innerhalb von Familien bestehen, wodurch die Ränge von Familien und Individuen und die Schätzwerte 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, SHIUE and PAULEY 1961, STONECYPHER 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 interfamily 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

<sup>1)</sup> Southern Forestry Research Center, P. O. Box 1060, Hot Springs, Arkansas 71901 U.S.A.

<sup>2)</sup> Weyerhaeuser Technology Center, Tacoma, Washington 98477 U.S.A.

<sup>3)</sup> Western Forestry Research Center, P. O. Box 420, Centralia, Washington 98531 U.S.A.

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

Genetic Test	Planting Location	Number of families	Family* Type	Number of blocks with row plots	Number of blocks with noncontiguous plots	Age-Years
1	Beaufort Co., N.C.	24	FS	4 (1 set)	4 (1 set)	5
2	Greene Co., Ala.	21	FS	8 (2 sets)	4 (1 set)	4
3	Kemper Co., Miss.	39	OP	8 (2 sets)	4 (1 set)	4
4	Greene Co., Ala.	46	FS	8 (2 sets)	4 (1 set)	4
5	Greene Co., Ala.	24	FS	4 (1 set)	4 (1 set)	4
6	Garland Co., AR	20	OP	8 (2 sets)	4 (1 set)	5
7	Sevier Co., AR	18	OP	8 (2 sets)	4 (1 set)	5
8	Howard Co., AR	17	OP	8 (2 sets)	4 (1 set)	5

\*FS = Full-sib, OP = Open-pollinated

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

Source	d.f.	Mean Square	f-test	Expected Mean Squares
Blocks	(b-1)	MS <sub>B</sub>	MS <sub>B</sub> /MS <sub>BxF</sub>	
Families	(f-1)	MS <sub>F</sub>	MS <sub>F</sub> /MS <sub>BxF</sub>	$\sigma_W^2 + n \sigma_{BxF}^2 + nb\sigma_F^2$
Block x Family	(b-1)(f-1)	MS <sub>BxF</sub>	MS <sub>BxF</sub> /MS <sub>W</sub>	$\sigma_W^2 + n \sigma_{BxF}^2$
Within Plots	$\sum(n_{ij}-1)$	MS <sub>W</sub>		$\sigma_W^2$

Where: b= number of blocks = 4  
 f= number of families  
 n<sub>ij</sub> = number of surviving trees/plot for the ith family in the jth block

Note: n will vary in the Expected Mean Squares depending on variable survival by plot, block, and family.

Table 3.1) — 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.

Test-Set	$\sigma_F^2$ (m)		$\sigma_{BxF}^2$ (m)	
	Plot Type		Plot Type	
	Row	Noncont.	Row	Noncont.
1-1	.045 *	.035 *	.003 NS	-.002 NS
2-1	.069 *	.064 *	.022 NS	-.020 NS
2	.037 *		.022 NS	
3-1	.077 *	.088 *	.074 *	-.002 NS
2	.073 *		.081 *	
4-1	.012 *	.025 *	.040 *	-.004 NS
2	.014 *		.087 *	
5-1	.048 *	.012 NS	.069 *	.005 NS
6-1	.012 NS	.015 *	.031 *	.014 NS
2	.007 NS		.024 *	
7-1	.017 *	.007 NS	.023 *	-.013 NS
2	.022 *		.015 *	
8-1	.015 NS	.012 *	.112 *	-.002 NS
2	.031 *		.020 NS	
Average of Tests	.036	.032	.043	-.003

1/  $\sigma_F^2$  = Family variance

$\sigma_{BxF}^2$  = Block-by-family interaction variance

\* Significant at the .05 probability level.

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

$$C.V. (\text{Plot Means}) = \frac{\sqrt{MS_{BxF}/n_H}}{\bar{X}} \quad 1.$$

Where: n<sub>H</sub> = Harmonic mean number of surviving trees per plot

$\bar{X}$  = Overall mean of the set

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  $\pm 5\%$  with a 5% 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.

Test-Set	Row Plots		Noncontiguous Plots	
	C.V.	b	C.V.	b
1-1	.065	9	.083	13
2-1	.109	20	.092	15
2	.108	20		
3-1	.112	21	.091	15
2	.109	20		
4-1	.102	18	.084	13
2	.084	13		
5-1	.123	25	.095	16
6-1	.065	9	.058	7
2	.061	8		
7-1	.098	17	.063	8
2	.094	16		
8-1	.089	14	.082	13
2	.136	30		
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  $\pm$  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:

$$b = \frac{t^2 (\text{C.V. Plot Means})^2}{d^2} \quad 2.$$

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 iteratively from the t-table and in-

termediate 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.

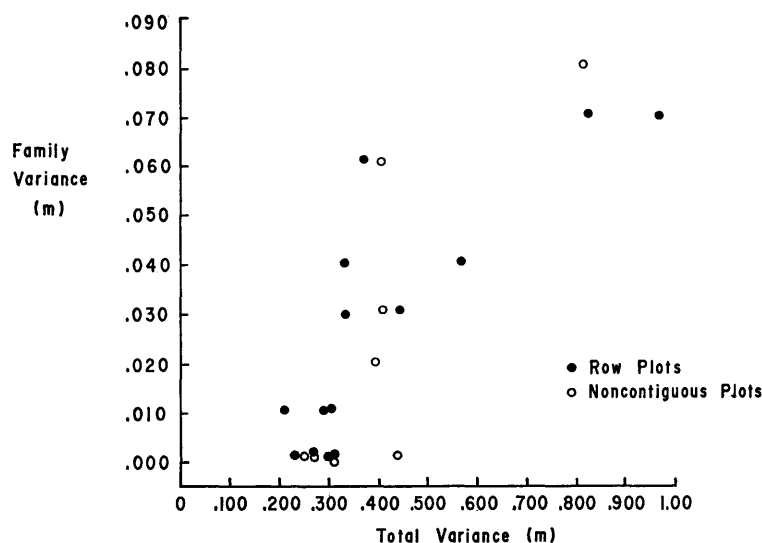


Figure 1. — Family variance component plotted against total variance for several data sets with row or noncontiguous plots. The data were from four- and five-year-old genetic tests of loblolly pine.

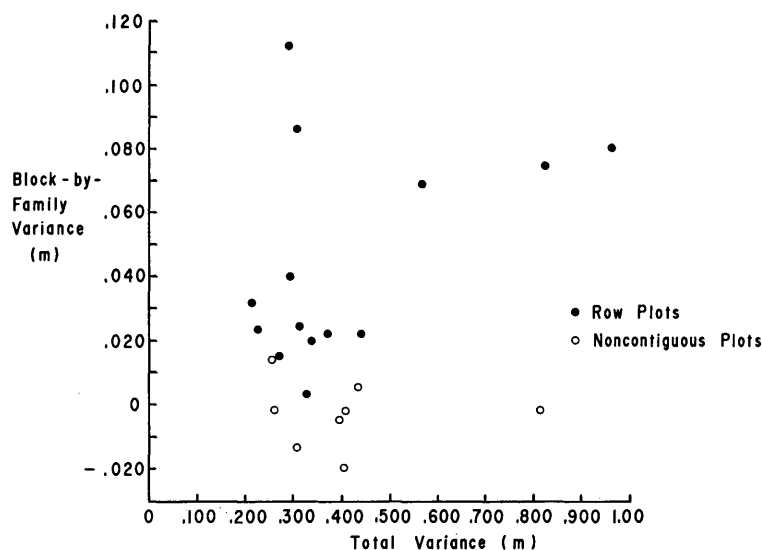


Figure 2. — Blocks-by-family interaction variance component plotted against total variance for several data sets with row or noncontiguous plots. The data are from four- and five-year-old genetic tests of loblolly pine.

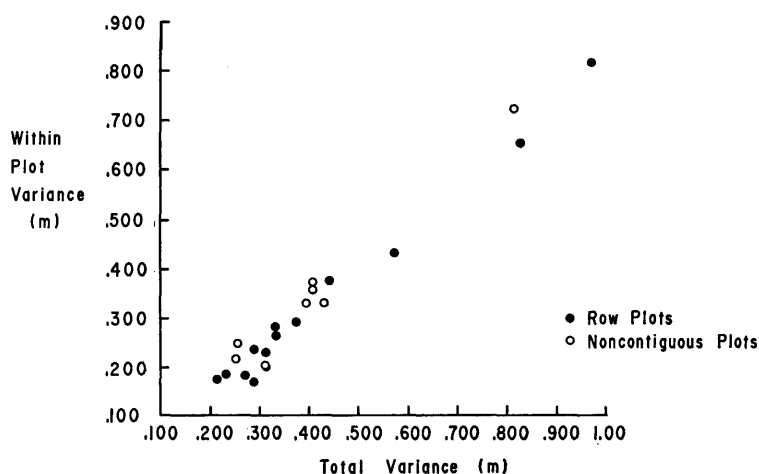


Figure 3. — Within plot variance plotted against total variance for several data sets with row or noncontiguous plots. The data are from four- and five-year-old genetic tests of loblolly pine.

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  $MS_{B \times F}$  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  $\times$  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 would 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 noncontiguous than for row plots — 60 versus 85 in an average

four-our five-year-old loblolly 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

CONKLE, M. T.: The determination of experimental plot size and shape in loblolly and slash pines. North Carolina State University, School of Forestry, Tech. Rep. No. 17, pp. (1963). — EVANS, T. C., J. C. BARBER and A. E. SQUILLACE: Some statistical aspects of progeny testing. Proc. Sixth Southern Conf. on Forest Tree Improvement, School of Forestry, University of Florida, Gainesville, June 7—8, pp. 73—79 (1961). — LIBBY, W. J. and C. C. COCKERHAM: Random noncontiguous plots in interlocking field layouts. *Silvae Genetica* 29 (5—6): 183—190 (1981). — SHIUE, C. J. and S. S. PAULEY: Some considerations on the statistical design for provenance and progeny tests in Tree improvement programs. *Forest Science* 7 (2): 116—122 (1961). — SNEDECOR, G. W. and W. G. COCHRAN: Statistical methods. The Iowa State University Press, Ames, Iowa. 593 pp. (1967). — STONECYPHER, R. W.: The loblolly pine heritability study. Tech. Bulletin No. 5. Woodlands Department Southlands Experimental Station, International Paper Company. 128 pp. (1966). — WRIGHT, J. W.: Introduction to forest genetics. Academic Press, Inc., New York, N.Y. 463 pp. (1976).

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

By P. P. COTTERILL<sup>1)</sup>, R. L. CORRELL<sup>2)</sup> and R. BOARDMAN<sup>3)</sup>

(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, genetic value, log transformation, least-squares.

#### Zusammenfassung

Es werden sechs mathematische Berechnungsarten beschrieben und bezüglich der relativen Genauigkeitsaussage des durchschnittlichen Wachstums von Kiefern, von denen ein Elternteil bekannt ist, miteinander verglichen. Die Daten sind jedoch keineswegs vollständig, da zwar einige Kiefernfamilien aufgeführt werden, aber nicht immer alle

entsprechenden Nachkommenschaftsprüfungen durchgeführt wurden. Hierzu dienten als Ausgangspunkt die aufgezeichneten Daten über das Stammvolumen als auch diejenigen der Geradschäftigkeit (eine sogenannte „Fünf-Punkte-Beurteilung“), die okular in fünf verschiedenen Nachkommenschaftsversuchen mit der Kiefernart *Pinus radiata* in Südaustralien vorgenommen wurde. Die Methoden waren: Rang/Punktzahlen RS, Standortanpassung SA, Standard-Standortanpassung SSA, verglichen mit den kleinsten Quadraten LS, den gewichteten kleinsten Quadraten WLS und den geschrumpften kleinsten Quadraten SLS. Zur Stabilisierung der Volumen-Standort Varianz wurde die logarithmische Umwandlung angewandt.

Dabei wurde festgestellt, daß die LS-, WLS- und SLS-Verfahren sehr weit bezüglich der Wesensmerkmale 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 Ergebnisse für die Ermittlung der Stammgeradheit der Familien.

#### 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  $\times$  progeny test array of available data is complete (all families represented at all sites) the analysis is straightforward in that family performance can be obtained by simply averaging across sites. More often, the family  $\times$  progeny test table is incomplete with a proportion of families represented at only a few sites, and in this case the non-orthogonality of the data can complicate

<sup>1)</sup> Division of Forest Research, CSIRO, Cunningham Laboratory, 306 Carmody Road, St. Lucia, Queensland 4067, Australia.

<sup>2)</sup> Division of Mathematics and Statistics, CSIRO, Private Bag No 2, Glen Osmond, South Australia 5064, Australia.

<sup>3)</sup> Woods and Forests Department, G. P. O. Box 1604, Adelaide, South Australia 5001, Australia.