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 silviculturally 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 Welsh who were instrumental in the establishment, maintenance and measurement of the tests, Barbara Bower who constructed the figures, David Poll who assisted in computer analyses and Cheryl 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, genet value, log transformation, least-squares.

Zusammenfassung

Dabei wurde festgestellt, daß die LS-, WLS- und SLS-Verfahren sehr weit bezüglich der Wachstumsmerkmale 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 × 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 × 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

1) Division of Forest Research, CSIRO, Cunningham Laboratory, 306 Carmody Road, St. Lucia, Queensland 4067, Australia.
2) Division of Mathematics and Statistics, CSIRO, Private Bag No 2, Glen Osmond, South Australia 5064, Australia.
3) Woods and Forests Department, G. P. O. Box 1604, Adelaide, South Australia 5001, Australia.
the estimation of average family performance. If the non-
orthogonality is slight, the problem may be avoided by
discarding data to give orthogonality (Nikles et al. 1977).
When the incompleteness is more substantial, average fa-
mily performance is sometimes expressed relative to other 'control' families which happen to occur at all sites, but the
success of this method depends directly on the stability of
the controls (Silvey 1978). The most common approach is
to use some type of mathematical procedure which adjusts
differences for differences between the group of sites at which a fa-
mily is represented and those where it is not.

In this study we have used stem volume and straightness
data from five open-pollinated progeny tests of Pinus ra-
diata D. Don in South Australia to compare six methods of
analysing incomplete progeny data. These methods are
identified as rank-score (RS), site-adjustment (SA), stan-
ard site-adjustment (SSA), least-squares (LS), and weight-
ted (WLS) and shrunken (SLS) least-squares. No allowance
was made for family \times\ progeny test interactions. RS, SA
and SSA are simple procedures which require no elaborate
computations and have been used in practical tree breed-
ing in Australia, but to our knowledge their accuracy has
not been previously tested. Least-squares are more reliable
methods which make simultaneous adjustment for incom-
pleteness of data in estimating both family and site ef-
effects, but for large data sets they require substantial
computing facilities. LS has been used to analyse incomple-
t progeny data in plant breeding (Patterson 1978, Silvey 1978,
Patterson and Silvey 1980), and WLS and SLS have been
used for the same purposes in animal breeding (Cunning-
ham 1965, Harvey 1968, Miller et al. 1968), although in an-
imal breeding the methods are used under different names
discussed later). No use of the methods of WLS or SLS has
apparently been made in forestry.

The aim of this study was to determine the accuracy of
the simpler procedures by comparing them with least-
squares. The advantage of using a logarithmic transforma-
tion in analysing growth data is also demonstrated.

**Progeny Data**

The open-pollinated progeny tests are identified as tests
5031, 5042, 5048, 5410 and 5411 (Australian Progeny Trial
numbers) and were established between 1969 and 1971
across five sites in the south-east of South Australia (as
detailed by Cotterill and Zed 1980). Appendix I shows the
highly incomplete family \times\ progeny test table. There are
65 families represented in one or more of the five progeny
tests but only one family (10954) occurs in all tests. The
sites can be regarded as representative of the environ-
mental conditions in the breeding region.

Each test has a randomised complete blocks design with
single-row plots (Cotterill and Zed 1980). The number of
block replications in each test is given in Table 1. The
height and diameter (at 1.3 m) of all trees in test 5031 were
measured at 10\% years after planting and at 71\% years for
the other tests. Stem volume was estimated from height
diameter by a simple conical function. Stem straight-
ness was assessed as a five-point visual score (1 = worst,
5 = best stem straightness).

**Mathematical Procedures**

The data were reduced to class means before analysis.
This does not waste information since records on indi-
vidual trees are not considered when calculating any stan-
dard errors of the effects of different families. Within a
test the families were randomly assigned to plots in ex-
perimental blocks, so the relevant error for comparing
families is the family \times block interaction. Calculation of
this experimental error involves only plot means. When
the effects of families are being estimated across sites the
family \times site interaction is the appropriate error for esti-
mating the standard errors of the relative performances
of the families.

**Log Transformation**

Differences in the age at assessment and environmental
conditions were reflected in large differences between-
sites for stem volume and the standard deviations of the
volume data varied almost proportionally to the means
(Table 1). This causes problems in assessing progeny test
data where the objective is to analyse the effects of fam-
ilies across sites. A correlation between the site means and
family effects would tend to make the family, site (and
block) effects multiplicative rather than additive. Under
these circumstances a log transformation is effective in
both stabilising the variance and achieving additivity.
In this instance we used base 10 logarithms but any base
would have been satisfactory. The effect of this transfor-
mation on the variances of the different sites is shown in
Table 1.

The means for straightness were relatively uniform across
sites (Table 1) and preliminary analyses showed that log transformation was not required. Analysis of only
raw straightness data is reported here.

**Rank-score**

This is a non-parametric method derived from the quart-
tile ranking used by Eldridge (1974). Families at a particu-
lar site are grouped into quartiles, and those in the upper
quartile are given a score +1, those in the middle half
are scored zero, and those in the lower quartile −1. The
mean score for each family is then found by averaging
these scores across sites. In this study, families at each site
(or test) were ranked in reverse order of merit and these
ranks were converted to scores by dividing by the number
of families at the site. The rank-scores for each family
were then averaged across all sites.

**Site-adjustment**

In this method an attempt to remove the effects of sites
was made by expressing the plot mean data as deviations
from the site means. The deviations for each family were
first averaged within-sites and then across-sites. The
method is equivalent to the herdmate-comparison used in
animal breeding (Miller et al. 1968).

**Standard Site-adjustment**

The problem of unequal variances across sites can be
reduced by dividing the previously mentioned deviations
from the site means by the site standard deviations (means
and standard deviations given in Table 1) thereby con-
verting the data to standard normal deviates. The stan-
dard normal deviates for each family were averaged with-
in-sites and then across-sites.

**Least-squares**

If the effect of the ith family is represented by \(a_i\) and
the jth block by \(b_j\), then within a site the model
\[
Y_{ij} = \mu + a_i + b_j + \epsilon_{ij}
\]
(I)
is fitted to the plot means \(Y_{ij}\) for each trait, choosing the
\(a_i\) and \(b_j\) constants to minimise the sum of squares of the
\(\epsilon_{ij}\)'s. This is straightforward as the families and blocks are
orthogonal within a site. To enable results to be generalised over a breeding region it is necessary to estimate family performance across a range of sites. If the effect of the kth site is designated \( t_k \), then across-sites the model

\[
Y_{ijk} = \mu + a_i + t_k + e_{ijk}
\]

(2)

is fitted to the family means \( \{Y_{ijk}\} \) for each trait. Usually the families and the sites will not be orthogonal necessitating some least-squares fitting of constants. The technique of fitting constants for non-orthogonal data was first considered by Yates (1935) and more recently by Patterson (1978). The \( a_i \) constants fitted for family effects in model (2) are unbiased, and if the progeny tests have equal precision, they will be minimum variance estimators with variance equal to the residual mean square divided by the number of sites across which a family occurs. For open-pollinated progeny tests the \( a_i \) constants for each family can be considered as least-squares estimates of general combining ability and \( 2a_i \) as least-squares estimates of breeding value.

There are many algorithms available for fitting least-squares constants. In this study it was simplest to remove the site effects with covariates and then proceed with a one-way analysis classified by families. Only a \( 4 \times 4 \) matrix is then inverted.

### Weighted Least-squares

One assumption made in the previous method of least-squares is that each progeny test has the same precision. This is not correct for the present data as the number of block replications varied from test to test (Table 1). If there are large differences in precision, more notice should be taken of the more precise tests. We have done this by weighting the data inversely proportional to the residual mean square for the site, which leads to a weighted least-squares analysis. Again it is only necessary to invert a \( 4 \times 4 \) matrix if a covariate analysis is used. The precision of each progeny test is obtained from a two-way analysis using model (1).

### Shrunken Least-squares

In this method families are considered to be distributed about some overall mean with variance \( \sigma^2_{a_i} \). This information is then combined with the experimental data to give estimates of family effects which are 'shrunk' towards the overall mean by an amount which depends on the effective replication of each family and \( \sigma^2_a \). We used the equation:

\[
a^*_i = \left[ \sigma^2_a (\sigma^2_e + \sigma^2_f) \right] a_i
\]

(3)

where \( a^*_i \) is the shrunken least-squares estimate of the effect of the ith family, \( a_i \) is the unweighted least-squares estimate obtained using model (2), and \( \sigma^2_a \) is the effective variance of the \( a_i \). Both \( \sigma^2_a \) and \( \sigma^2_e \) were obtained from the unweighted least-squares analysis using the method given by Hudson (1966). In the case of equation (3) the \( a^*_i \) are shrunk towards zero which is the overall mean of the \( a_i \). The \( a^*_i \) estimates are more conservative for the purpose of family recommendations in the sense that poorly replicated families (with a larger \( \sigma^2_e \)) are shrunk more towards zero than those which are better replicated.

The method of shrunken least-squares has been outlined by Lindley and Smith (1972) and is similar to the methods of regressed least-squares used by Harvey (1968), and weighted least-squares used by Cunningham (1965) and Miller et al. (1968). We refer to the methods as shrunken least-squares to avoid the ambiguity of the word 'regressed' and confusion with the method of weighted least-squares used in this study.

### Results and Discussion

Ordinary analysis of variance of data using model (2) revealed that both families and sites had a significant effect on all traits studied, but that sites had an overwhelming effect on volume (Table 2). The magnitude of the effects of sites on volume and the corresponding reduced relative contribution of families to the overall sums of squares makes this growth trait a stringent evaluation of the efficiency and reliability of the various methods in removing the variation due to sites and extracting information on families from the incomplete data. To take the opposite extreme, when there are only small site effects the relative performance of families could be reliably deduced by simply averaging the incomplete data across sites. In the case of straightness, site effects made a much smaller contribution to the overall sums of squares and this trait therefore provides a less stringent evaluation of the various methods. This is evident from subsequent results.

Assumptions made by the various methods of analysing incomplete data are summarised in Table 3.

---

**Table 1.** Number of block replications in the randomised blocks design of each progeny test, and the means and standard deviations (s.d.) for stem volume, stem volume after transformation of data using base 10 logarithms, and stem straightness.

<table>
<thead>
<tr>
<th>Proleny test</th>
<th>5031</th>
<th>5032</th>
<th>5038</th>
<th>5410</th>
<th>5411</th>
</tr>
</thead>
<tbody>
<tr>
<td>Replications</td>
<td>12</td>
<td>6</td>
<td>14</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Volume (dm³)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>165</td>
<td>95</td>
<td>76</td>
<td>86</td>
<td>45</td>
</tr>
<tr>
<td>s.d.</td>
<td>34.2</td>
<td>10.0</td>
<td>22.7</td>
<td>23.8</td>
<td>12.3</td>
</tr>
<tr>
<td>Volume after transformation (log_{10} dm³)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>2.21</td>
<td>1.99</td>
<td>1.88</td>
<td>1.95</td>
<td>1.64</td>
</tr>
<tr>
<td>s.d.</td>
<td>0.093</td>
<td>0.084</td>
<td>0.076</td>
<td>0.069</td>
<td>0.098</td>
</tr>
<tr>
<td>Straightness (1 to 5-point visual score)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>3.04</td>
<td>2.79</td>
<td>2.80</td>
<td>2.55</td>
<td>2.48</td>
</tr>
<tr>
<td>s.d.</td>
<td>0.508</td>
<td>0.365</td>
<td>0.326</td>
<td>0.316</td>
<td>0.361</td>
</tr>
</tbody>
</table>

**Table 2.** Mean-squares and, in brackets, F-ratios from unweighted least-squares analysis of progeny data using model (2). Volume is after transformation of data using base 10 logarithms.

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>Mean squares</th>
<th>Stem straightness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family</td>
<td>64</td>
<td>0.0032</td>
<td>(2.1)</td>
</tr>
<tr>
<td>Site</td>
<td>4</td>
<td>0.7227</td>
<td>(48.1)</td>
</tr>
<tr>
<td>Residual</td>
<td>145</td>
<td>0.0015</td>
<td>(57.1)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>Mean squares</th>
<th>Volume straightness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family</td>
<td>64</td>
<td>0.0032</td>
<td>(2.1)</td>
</tr>
<tr>
<td>Site</td>
<td>4</td>
<td>0.7227</td>
<td>(48.1)</td>
</tr>
<tr>
<td>Residual</td>
<td>145</td>
<td>0.0015</td>
<td>(57.1)</td>
</tr>
</tbody>
</table>
only that there is a random distribution of families across sites (or expressed another way, that the 'quality' of the families at each site in compatible) and is therefore quite a robust procedure. WLS is also robust assuming only additivity of family and site effects, as implicit in model (2). Even when non-additivity is present in the data, least-squares analysis of cell means (or in this case family means) using model (2) should still provide reliable comparisons between treatments (Urbash and Weeks 1978). SA can be considered an approximation to the fitting of constants by LS but requires the assumption of random distribution of families across sites.

Of the six methods we prefer WLS due to its providing minimum variance unbiased estimates of the effects of families under the additive model, and we have used WLS to evaluate the relative accuracy of the other methods. Since the purpose of progeny testing is to rank families in order of merit, the association between the ranking according to WLS versus ranking according to the other methods has been used to judge 'relative accuracy'.

Despite theoretical differences in complexity the six methods were found to rank families in about the same order for volume and almost exactly the same order for straightness. The actual relationships between rankings according to the method of WLS and rankings according to each of the other methods are quantified in Table 4 by Spearman's rank correlations. A rank correlation of 1.0 indicates complete agreement in the order of the ranks while zero indicates complete disagreement. Table 4 also gives, for each trait, the number of the top 10 families as selected by WLS which would have been selected by each of the other methods. The top 10 families are used here because that happens to be the number of families actually retained from these combined progeny tests in breeding *P. radiata* in South Australia.

Each of the least-squares procedures gave essentially the same rankings of families (rank correlations between WLS and LS or SLS were 0.97 for volume and 0.99 for straightness; Table 4) and little real advantage seems to have been achieved by using the slightly more complicated SLS or WLS procedures compared to unweighted LS. This finding agrees with that of Miller et al. (1968) who reported rank correlations of 0.99 between LS and SLS rankings of dairy sires determined from incomplete progeny data. For both volume and straightness LS and SLS selected either nine or all of the 10 families chosen by WLS (Table 4). It was noticeable that SLS is more conservative in recommending families which are not well replicated. For instance, WLS ranked families 80055, 50078 and 50047 first, second and third respectively, for volume, but because 800055 was represented at only two sites (Appendix 1) SLS reduced it to rank third behind 50078 and 50047 which were both represented at three sites.

It may be unreasonable to discard a promising variety because little is known about it, as could occur with SLS. On the other hand Patterson and Silvey (1980) have argued that a family is more likely to be chosen by LS if its test means exceed its true mean. For the purpose of family recommendations and estimating responses from selection the more conservative SLS estimates may be preferred because they remove this potential bias of LS. The correction is however dependent on having good estimates of the various components $\sigma^2$, and $\sigma^2$K for use in equation (3).

The methods of RS and SSA showed comparable accuracy (rank correlations with WLS were 0.89 to 0.91 for volume and 0.99 for straightness; Table 4) while SA was marginally less reliable for volume (rank correlation with WLS equalled 0.82). In analyses not reported here we have found the accuracy of SA is further reduced when the method is used to analyse raw volume data not transformed logarithmically, because unlike RS and SSA, the method of SA requires the assumption of uniform variance across sites (Table 3). Nevertheless, RS, SA and SSA all selected
Appendix I. — Representation of families across progeny tests.

<table>
<thead>
<tr>
<th>Family</th>
<th>Progeny test</th>
<th>Family</th>
<th>Progeny test</th>
</tr>
</thead>
<tbody>
<tr>
<td>30040</td>
<td>X</td>
<td>50267</td>
<td>X</td>
</tr>
<tr>
<td>50001</td>
<td>X</td>
<td>50268</td>
<td>X</td>
</tr>
<tr>
<td>50006</td>
<td>X</td>
<td>50269</td>
<td>X</td>
</tr>
<tr>
<td>50009</td>
<td>X</td>
<td>50055</td>
<td>X</td>
</tr>
<tr>
<td>50010</td>
<td>X</td>
<td>10935</td>
<td>X</td>
</tr>
<tr>
<td>50012</td>
<td>X</td>
<td>10948</td>
<td>X</td>
</tr>
<tr>
<td>50013</td>
<td>X</td>
<td>10994</td>
<td>X</td>
</tr>
<tr>
<td>50015</td>
<td>X</td>
<td>10996</td>
<td>X</td>
</tr>
<tr>
<td>50016</td>
<td>X</td>
<td>10967</td>
<td>X</td>
</tr>
<tr>
<td>50017</td>
<td>X</td>
<td>10984</td>
<td>X</td>
</tr>
<tr>
<td>50018</td>
<td>X</td>
<td>12061</td>
<td>X</td>
</tr>
<tr>
<td>50022</td>
<td>X</td>
<td>12038</td>
<td>X</td>
</tr>
<tr>
<td>50024</td>
<td>X</td>
<td>12060</td>
<td>X</td>
</tr>
<tr>
<td>50028</td>
<td>X</td>
<td>12112</td>
<td>X</td>
</tr>
<tr>
<td>50030</td>
<td>X</td>
<td>12130</td>
<td>X</td>
</tr>
<tr>
<td>50031</td>
<td>X</td>
<td>12187</td>
<td>X</td>
</tr>
<tr>
<td>50039</td>
<td>X</td>
<td>12197</td>
<td>X</td>
</tr>
<tr>
<td>50042</td>
<td>X</td>
<td>12236</td>
<td>X</td>
</tr>
<tr>
<td>50043</td>
<td>X</td>
<td>12247</td>
<td>X</td>
</tr>
<tr>
<td>50044</td>
<td>X</td>
<td>12294</td>
<td>X</td>
</tr>
<tr>
<td>50045</td>
<td>X</td>
<td>12315</td>
<td>X</td>
</tr>
<tr>
<td>50047</td>
<td>X</td>
<td>12349</td>
<td>X</td>
</tr>
<tr>
<td>50048</td>
<td>X</td>
<td>12251</td>
<td>X</td>
</tr>
<tr>
<td>50077</td>
<td>X</td>
<td>12373</td>
<td>X</td>
</tr>
<tr>
<td>50078</td>
<td>X</td>
<td>12274</td>
<td>X</td>
</tr>
<tr>
<td>50079</td>
<td>X</td>
<td>12378</td>
<td>X</td>
</tr>
<tr>
<td>50080</td>
<td>X</td>
<td>12403</td>
<td>X</td>
</tr>
<tr>
<td>50082</td>
<td>X</td>
<td>12408</td>
<td>X</td>
</tr>
<tr>
<td>50126</td>
<td>X</td>
<td>12412</td>
<td>X</td>
</tr>
<tr>
<td>50127</td>
<td>X</td>
<td>12419</td>
<td>X</td>
</tr>
<tr>
<td>50176</td>
<td>X</td>
<td>12423</td>
<td>X</td>
</tr>
<tr>
<td>50177</td>
<td>X</td>
<td>12447</td>
<td>X</td>
</tr>
<tr>
<td>50178</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Families identified by Australian Pine Tree Register numbers.

The data will also affect the relative accuracy of alternative methods. Differences between the simpler methods and least-squares would be expected to be less when the family × progeny test Table is more complete.

We could argue that the amount of effort required in the field to establish and measure progeny tests is sufficient justification for carrying out the additional computations required by the least-squares procedures. This is in view of the fact that most breeders have access to suitable computing facilities to calculate the least-squares estimates.

The six methods can also be applied to control-pollinated progeny data, although the simpler methods would be unsatisfactory where there is incompleteness in both the crossing scheme and the representation of full-sib families across sites. If there were substantial incompleteness across both crossing scheme and sites then even the least-squares estimates would not be very reliable (USQUHART and WEEKS 1978). GILBERT (1967) described the application of an unweighted LS approach for analysing progeny data from incomplete crossing schemes, and HARVEY (1960 and 1968) outlined the application of LS and SLS procedures using the more complex models needed to analyse full-sib progeny data from a range of sites.

Acknowledgements

Thanks are given to staff of the Woods and Forests Department, CSIRO and Softwood Holdings Ltd who collected the data. It is a pleasure to acknowledge the initiative of CHARLES FAWCETT who established some of the tests reported here.

Literature Cited