PROGENY TEST DESIGN AND ANALYSIS

J. P. van Buijtenen

Abstract.--A workshop on progeny testing was held in Auburn, Alabama on June 15-16, 1982, sponsored by Cooperative Regional Project S-23, "Breeding Strategies for Genetic Improvement of Commercial Forest Trees in the South".

Topics covered were objectives of progeny tests; mating designs; raising progeny test seedlings; field designs; test establishment, maintenance, and measurement; and data handling and analysis.

This paper summarizes parts of the workshop, particularly the topics related to the design and analysis of progeny tests. Emphasis is placed on the practical rather than theoretical aspects, although both were covered in the workshop.

Copies of the proceedings can be obtained from any of the cooperating state agricultural experiment stations.

Additional keywords: Polycross, diallel, factorial design, nested design, field layout, data handling.

INTRODUCTION

The following paper is a summary of the highlights of the S-23 Workshop on progeny testing held in Auburn on June 15-16, 1982, emphasizing the design and analysis aspects. It will touch on mating designs, field layout, and the management and analysis of progeny test data. The choice of designs depends to a large extent on the objectives and the overall breeding strategy of the particular tree improvement program involved. As pointed out by McKinley (1983), the most common objectives of progeny tests are: a) to provide information for evaluating parents; b) to estimate genetic parameters; c) to produce a base population for advanced generation selection; and d) to estimate realized gain directly.

Since it is not always possible to meet the progeny test objectives with a single progeny test, it is often necessary to assign priorities. Such priorities are sometimes difficult to identify and may depend on the goals of the organization involved, the timing of the progeny test and the generation one is working on. In the inception of a breeding program, delineation of seed sources may rank very high, while at a later stage advanced generation breeding may be the primary goal.

1/ Principal Geneticist, Texas Forest Service and Professor, Texas Agricultural Experiment Station, College Station, Texas.

MATING DESIGNS

There is a large variety of mating designs. They can be divided into two main groups, single designs and complementary designs. Since in a practical breeding program, one usually tries to achieve several objectives at one any single design may be reasonably useful for all purposes, but not necess ily the best for any one of them. The concept of complementary designs is based on the philosophy that it can be an advantage to combine several simple designs, each of which is very efficient for one particular purpose.

Open-Pollinated Designs

Open-pollinated progeny tests may be obtained from ortets as they occur in the woods or ramets in seed orchards. Selections from open-pollinated progeny tests obtained from seed orchards may be related, and crosses among them may involve inbreeding and all the problems involved with inbreeding depression. Using open-pollinated seed from ortets in the woods avoids this problem, but one loses the gain associated with selection of the male parent. Open-pollinated tests are useful to evaluate parents if the pollen is a good mixture and does not vary much from tree to tree, but may be unreliable if each female parent has a unique and different pollen sample. The experience of several organizations indicates that there may indeed be a problem, since the results of open-pollinated tests do not always agree with those of subsequently obtained control-pollinated tests. Open-pollinated tests of seed obtained out of a seed orchard would seem a very appropriate means of determining realized gain in that particular orchard.

The Polycross

The polycross design is in many ways similar to the open-pollinated test. A set of pollens is mixed and used to make control pollinations. Polycross tests are very good for determining general combining ability and, thus, for evaluating parents, but generally are not suitable for making advanced generation selections, since inbreeding may result because of relatedness between the offspring. Several interesting modifications of the polycross are discussed by van Buijtenen and Namkoong (1983).

The Complete Diallel

The complete diallel consists of all possible crosses among the available parents. This design obviously will give the maximum amount of information possible on the parents included in the test. The number of crosses, however, goes up as the square of the number of parents and the cost and time required very quickly become unmanageable. For an orchard including 40 clones, it would, e.g., take 1,600 crosses to make a complete diallel.

The Modified Half-Diallel

This is similar to the complete diallel, but omits the selfs and the reciprocal crosses. Although the number of crosses is somewhat less than half of that of the complete diallel, it is still quite large. Although it is still possible to estimate both gca and sca, it is no longer possible to study the effect of selfing and of making reciprocal crosses.

The Disconnected Half-Diallel

This is a variant of the modified half-diallel created by dividing the parents into equal sets and making a modified half-diallel within each set. A disconnected six-tree half-diallel, e.g., would involve 15 crosses per set. This kind of design is a reasonably good compromise that allows one to determine gca and sca. It is good not only for evaluating parents, but also for identifying outstanding specific crosses. It also is quite useful as a means of generating a population in which to make selections for the next generation. There is a problem, however, in that usually not all disconnected diallels can be planted in one field test, which means that one cannot capture the genetic variation between the sets. Particularly in small sets this can be a substantial amount since van Buijtenen and Burdon (1983) found that the amount of additive genetic variation between sets is equal to the total additive genetic variation divided by the number of parents. In the six-tree disconnected diallels this would leave, for instance, 1/6 of the additive genetic variation between the sets.

Partial Diallels

In a sense, any set of crosses short of a complete diallel could be called a partial diallel. However, one usually refers to a design consisting of one or more sets of crosses following the diagonals across the diallel tables. This again is a fairly good compromise that allows one to evaluate parents, estimate genetic parameters, and make selections for the next generation.

Factorial Designs

In a factorial design, two separate groups of parents are crossed with each other in all possible combinations. In the disconnected factorial, the same principle is followed as in the disconnected diallel: the population is sub-divided into groups and the trees within each group are crossed according to a factorial design. Pepper and Namkoong (1978) made a detail economic evaluation of both nested and factorial designs for progeny testing. The factorial design is generally slightly superior to the nested design in terms of evaluating parents. For evaluating genetic parameters, the factorial is usually not quite as efficient as the various diallel designs. For selection of new parents for the next generation, the factorial design can be quite satisfactory if there are enough male parents. The four tester design employed formerly was not very well suited, because of the high degree of relatedness among second generation selections which resulted from this design.

Nested Design

In a nested design, one parent is crossed with a number of other parents. Usually the male is the rare parent which is crossed with several females. Each female, however, is crossed with only one male. The nested design is not too efficient for many purposes, such as evaluating parents and estimating genetic parameters. It could, however, work very well in a complementary design where family selection is handled with a design such as a polycross, while the nested design provides a source of trees for the following generation. This would be particularly true if some of the parents are known already, in which case the best general combiners could be assigned to the rare parent.

FIELD DESIGNS

Field designs were discussed by Bridgwater et al. (1983). The bas is again that the field design needs to be appropriate for the particul jective being served. In an operational tree improvement program, quit a compromise needs to be struck between a number of practical considera and the needs of the various objectives to be accomplished. From the p tioner's point of view, the question to be answered is a very simple one: Where do I plant each tree? This is about where the simplicity ends. Bor ing a concept from computer science, we can take the top-down or bottom-up proach in discussing this problem and I'd like to follow the top-down appr

How Many Locations are Needed?

Philosophies in this respect seem to differ widely by region. In the South, the common procedure is to plant the same families on three or four different locations. In some western programs, families may be planted in to 16 locations. Which is the better approach depends on the type of tree wants to select and on the amount of genotype x environment interaction that is present.

Let's start by looking at the genotype x environment interaction. If no genotype x environment interaction is present, in other words if all families rank the same on all sites, one needs to test on only one site. If genotype x environment interaction is present, however, one needs to make a decision. One can either select a general purpose tree that may not grow the best at an) given location, but will grow reasonably well at almost all locations, or one can select genotypes that are specifically adapted to a given site. This will greatly influence the strategy of the testing.

If one wants to select for a generally adaptable type, one could plant relatively small numbers of trees per family in a given location and plant families on many locations. This should give one the best possible evalua although one would have only very limited information on the suitability ofa particular family for a given site.

If one wants to know something about the adaptability to a specific site, one needs to increase the number of trees per family planted at a given location. Economics, as a consequence, will force a reduction in number of test sites to keep the cost of progeny testing within reason.

Although I never heard anybody address this issue, I feel that the weste tree improvement programs lean towards the general adaptability approach, whil the southern programs are somewhat of a compromise evaluating general adaptab lity, but also obtaining some information on the performance of a family at a given site.

How Large Should a Test Be at a Given Location?

This is really not a meaningful question, since the size of the test itself is not critical. It is the size of a single replication that is critical

To maintain soil uniformity within a replication, it is necessary to keep the replication size as small as possible, about 1/4 to 1/2 acre maximum. There is no definite limit to the number of replications one could include. The number of replications needed depends on the precision desired and the plot size, which will be discussed later. A typical test includes about 2,000 trees, although I've seen tests of less than 500 trees and tests of around 10,000 trees.

How Many Families Should Be Included in a Test?

This again is closely tied to the size of the replications and the size of the plots. If we assume that the maximum replication size is about 1/4 acre and the spacing is 8 x 8 feet, then the relation between plot size and the maximum number of families included in a test is as depicted in Figure 1. For a 1/2 acre replication and 10-tree plots, this would amount to about 34 families. Using single tree plots, one could accomodate well over 100 families and keep the replications smaller than 1/2 acre.

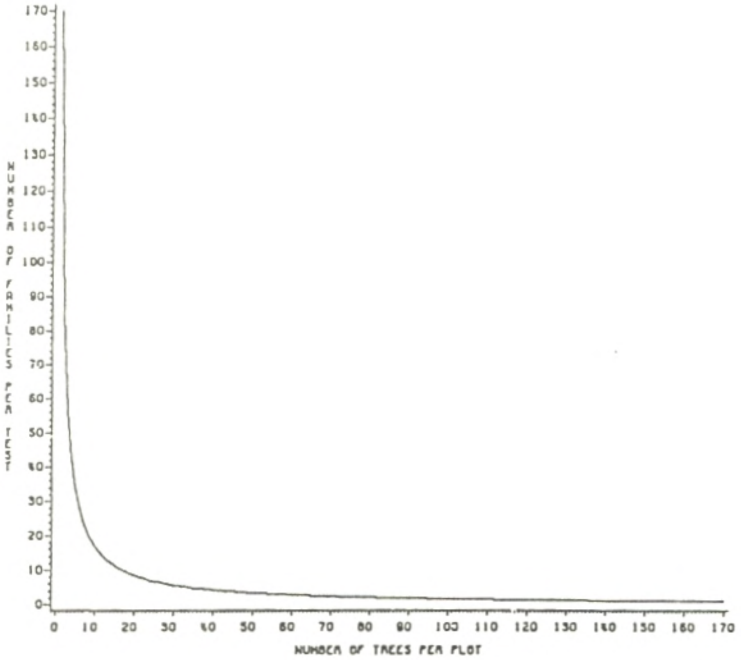

Figure 1. Relation between plot size and number of families per test, assuming a replication size of 1/4 acre and a spacing of 8 x 8 feet.

How Many Replications Per Location are Needed?

Again this question cannot be answered categorically, it depends on the precision desired and the plot size. There are sophisticated statistical formulae available to calculate this, but in practice the total number of trees needed per family remains relatively constant and is mainly distributed differently between replications and number of trees per plot. This is not quite true, however. Although one needs more replications with smaller plots, the total number of trees needed is somewhat less. In general, smaller plots and more replications will give greater precision. In our experience, six replications and four trees per plot are definitely not enough, while ten replications of 10-tree row plots are probably more than needed.

How Many Trees Per Plot are Needed?

This has been largely covered under previous headings, but a few things remain to be said. Over the last 30 years, experiments have been put in, which range from large 100-tree plots to single-tree plots. From a statistical point of view, the single-tree plots and a modification often called non-contiguous plots offer the greatest advantages. However, they are very difficult to establish and maintain. Hundred-tree plots are too large, since one could include only about six families in a 1/2 acre replication. As a consequence, there usually is a large family by replication interaction making it very difficult to distinguish between the performance of individual families. Small row-plots utilizing anywhere from 4 to 10 trees per plot currently seem to be most common. They allow a reasonable number of families per test, are easy to handle in the field, and give good statistical precision if a sufficient number of replications is used. They also lend themselves well to making second generation selections, as selected trees can be compared to the other trees within the same plot. Single-tree plots are very difficult to use for second generation selections, since members of the same families are scattered throughout a replication. We are currently looking at different ways of evaluating single-tree plots and non-contiguous plots, but a satisfactory solution has not been found yet.

Laying Out the Test in the Field

This is by far the most critical and most difficult phase of progeny test design. There is no such thing as a perfectly uniform piece of land and care must be taken to properly cope with the different kinds of site variability encountered. Following are some examples of site variability and what to do about it.

1) Slope.--The common way to reduce the impact of slope on the evaluation of the families is to lay out the replications parallel to the contour lines and the rows perpendicular. This way there may be some increased variability within the plots, but each family will sample roughly the same type of environment.

2) Pimple mounds, small depressions, and potholes.--Two approaches are possible here. If the area is too large, it is best not to include it in the test. This may cause the replications to be separated from each other, however, this is far preferable over including the soil variation. If the unusable areas are small, a common approach is to plant filler trees. These serve to reduce border effects and are not included in the measurements.

3) Distinct changes in soil type.--Try to keep one replication in the same soil type, thus, avoiding a change in soil type in the middle of one replication. Again, this may cause one replication to be separated or off-set from the others.

PROGENY TEST DATA HANDLING AND ANALYSIS

This was discussed in considerable detail at the S-23 meeting (Lowe et al. 1983), including the statistical aspects, but I will mostly highlight some of the practical aspects.

Data Collection

The most common way to do this is probably still by handwriting the numbers either on special forms or on 80-column computer coding forms. Regardless of the system used, it is nice to have the previous measurement available on the form to help identify position in the field, to record recent mortality, and use old measurements for confirming the validity of the new ones.

Recently new equipment has been introduced where measurements can be keyed immediately into an electronic recording device. In principle, it is possible to use a wand-like device to record tree identification just like it is done in grocery stores. Labels, however, are subjected to a lot of dirt, wear and tear in the field, and this approach has not turned out too well. The recording devices themselves, however, have been quite satisfactory, and a number of them are in use. Data can be transmitted over a telephone line to the main computer after which data can be edited and stored. It is also possible to load a file of previous measurements into the recording device which can be displayed while new data are recorded, giving the same advantages as recording new data along with previous measurements on paper. This field is developing so rapidly that before very long there will be several cheap and sophiscated devices on the market that will make a complete shift to automated data recording almost unavoidable.

Data Editing

In addition to proofreading, it is advantageous to run the data through an editing program that will flag potential errors. This can be done, for instance, by having it compare new measurements to previous measurements, and by specifying ranges within which the data are expected to fall. Another way to flag potential errors is to calculate standard deviations for each plot and double check the plots with unusually large standard deviations. Other items that the editing program should check are the genetic identification codes, the number of replications, and the number of observations within each replication.

Data Storage and Backup

The data are usually stored on cards, tape, discs, or diskettes. It is absolutely essential to have some form of backup. In other words if the data are on tape, they should be on two separate tapes kept in two separate locations, or if the data is stored on a disc or diskette, they should be backed up by tape or loaded on a separate disk. Data stored on electro-magnetic devices are particularly vulnerable to complete or partial loss due to computer failures. In addition, it is a good idea to have a paper backup such as a recording sheet or a printout.

Statistical Considerations

There is no need to go into detail on the statistical aspects, but I would like to relate some of the major areas that were covered at the meet A major portion was devoted to the analysis of single location tests. Thi relatively straightforward if one has a well-balanced set of data in hand. These seem to be a rarity, however. Fortunately, a number of computer programs are available that handle unbalanced data relatively well, although loss of precision usually occurs.

A much more difficult problem is the analysis of tests planted across tiple locations. Inevitably some families will rank differently on one lit tion than on a second location. This could be due either to inaccuracies in the progeny test or because a particular family may be better adapted to one site than the other. As discussed earlier, at this point one needs to decide' whether one wants to select for a family adapted to a particular site or for families that are adaptable to a range of sites. Regardless of the decision,- one will have to determine how a given family responds to a range of site qualities. This can be done fairly readily using a regression approach as is illustrated in Figures 2 and 3. Some well-known seed sources are good examples of these type of responses. Livingston Parish loblolly pine, for ins has a rather steep response curve, indicating it does relatively poorly on site index land, but does extremely well where the site index is high. By contrast, the drought hardy sources from the Lost Pines area in Texas have a relatively flat response and do reasonably well on poor sites, but do not increase their growth as much as Livingston Parish when planted on good sites.

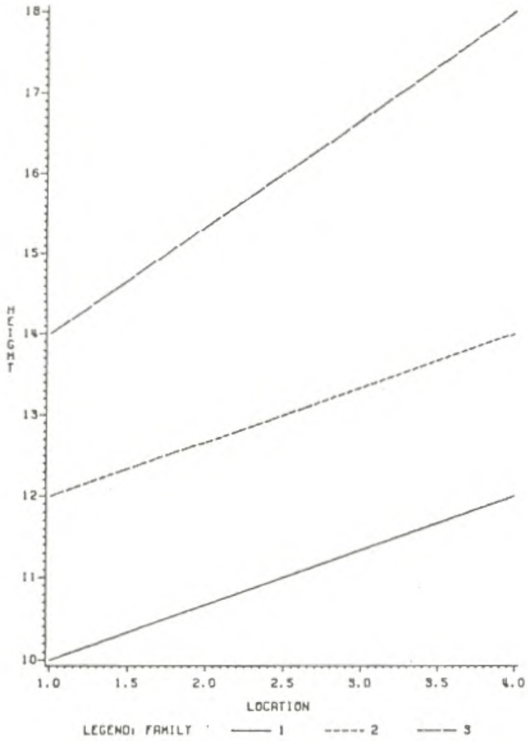

Figure 2. A genotype x environment interaction caused by an increase in variance among families on better sites.

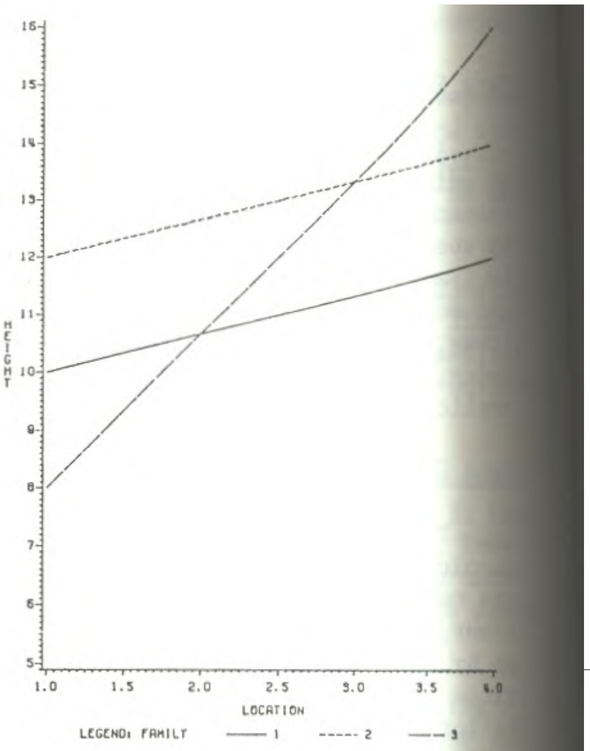

Figure 3. A genotype x environment interaction caused by a change in family rank at different locations.

Data Summarization Across Tests

This is an extension of the problem discussed in the previous section, except here the purpose is to rank families for roguing a seed orchard. Several major problems need solving. The data are usually unbalanced and the trees are planted over a range of sites. To put the data on a common basis, the family averages at a given location need to be expressed in a way that is independent of site. This can be done by expressing them as a percentage of the checks. This does not always work too satisfactorily because different tests may have different checks included, or they may have so few checks that the check doesn't form a stable base line. Commonly it does work well to express a family as the percent of the average of the progeny test. This in itself can create problems, because some progeny tests may be composed of better genetic material than others. Another good approach is to express the ranking of a family in terms of standard deviations. The N.C. State performance index is based on this principle (Hatcher et al. 1981).

A system that, as far as I know, has not been implemented anywhere, but which theoretically looks the best is to weight family averages by plantation according to their standard errors. In other words, the more reliable the observation, the higher weight it will carry. These weighted averages are then fitted across progeny test sites using a least squares procedure. This gets very demanding in terms of the amount of computer memory required, and the data needs to be complete enough where the values of the families can be properly estimated. No system is totally satisfactory yet, but good progress has been made.

CONCLUSIONS

The mating design and progeny test design need to be matched to the progeny test objectives. One has the choice between some reasonable compromise solutions and a complementary design consisting of a combination of single tests, each serving a specific objective.

A revolution in data collection is underway with the advent of electronic data recording devices. As fast as this field is developing, it seems likely that the clipboard and notepad may suffer the same fate as the slide rule.

Satisfactory means of analyzing each type of design are available, including programs that will handle unbalanced data. Several workable systems of summarizing data across tests are available, but there is a need for further improvements in technology in this area.

Data Summarization Across Tests

This is an extension of the problem discussed in the previous section, except here the purpose is to rank families for roguing a seed orchard. Several major problems need solving. The data are usually unbalanced and the trees are planted over a range of sites. To put the data on a common basis, the family averages at a given location need to be expressed in a way that is independent of site. This can be done by expressing them as a percentage of the checks. This does not always work too satisfactorily because different tests may have different checks included, or they may have so few checks that the check doesn't form a stable base line. Commonly it does work well to express a family as the percent of the average of the progeny test. This in itself can create problems, because some progeny tests may be composed of better genetic material than others. Another good approach is to express the ranking of a family in terms of standard deviations. The N.C, State performance index is based on this principle (Hatcher et al. 1981).

A system that, as far as I know, has not been implemented anywhere, but which theoretically looks the best is to weight family averages by plantation according to their standard errors. In other words, the more reliable the observation, the higher weight it will carry, These weighted averages are then fitted across progeny test sites using a least squares procedure. This gets very demanding in terms of the amount of computer memory required, and the data needs to be complete enough where the values of the families can be properly estimated. No system is totally satisfactory yet, but good progress has been made.

CONCLUSIONS

The mating design and progeny test design need to be matched to the progeny test objectives. One has the choice between some reasonable compromise solutions and a complementary design consisting of a combination of single tests, each serving a specific objective.

A revolution in data collection is underway with the advent of electronic data recording devices. As fast as this field is developing, it seems likely that the clipboard and notepad may suffer the same fate as the slide rule.

Satisfactory means of analyzing each type of design are available, including programs that will handle unbalanced data. Several workable systems of summarizing data across tests are available, but there is a need for further improvements in technology in this area.

LITERATURE CITED

Bridgwater, F, E., J. T. Talbert, and D. L. Rockwood. 1983, Field Design for Genetic Tests of Forest Trees. In Progeny Testing of Forest Trees, Proc. of S-23 Workshop on Progeny Testing, June 15-16, 1982, Auburn, Alabama, Southern Cooperative Series Bulletin No. 275, pp. 28-38.

Hatcher, A. V., F. E. Bridgwater, and R. J. Weir. 1981. Performance Level—Standardized Score for Progeny Test Performance. Silvae Genetica 30(6): 184-187.

Lowe, W. J., R. Stonecypher, and A. V. Hatcher. 1983. Progeny Test Data Handling and Analysis. In Progeny Testing of Forest Trees, Proc. of S-- Workshop on Progeny Testing, June 15-16, 1982, Auburn, Alabama, Southern Cooperative Series Bulletin No. 275, pp. 51-67.

McKinley, C. R. 1983. Objectives of Progeny Tests. In Progeny Testing of Forest Trees, Proc. of S-23 Workshop on Progeny Testing, June 15-16, 1982, Auburn, Alabama, Southern Cooperative Series Bulletin No. 275, pp. 2-5.

Pepper, W. D. and G, Namkoong. 1978. Comparing efficiency of balanced mating Designs for Progeny Testing. Silvae Genetica 27(3-4):161-169.

van Buijtenen, J. P. and G. Namkoong, 1983. Mating Designs. In Progeny Testing of Forest Trees, Proc. of S-23 Workshop on Progeny Testing, June 15 16, 1982, Auburn, Alabama, Southern Cooperative Series Bulletin No. 275, pp. 7-13.

van Buijtenen, J. P. and R. D. Burdon. 1983. Efficiency of Mating Designs for Second Generation Selection. (In Preparation).