COMPARISONS AMONG COMBINATIONS OF PROGENY

TEST DESIGNS AND SELECTION SYSTEMS FOR

MAXIMUM GENETIC GAIN IN LOBLOLLY

PINE (PINUS TAEDA L.)

By

JUDITH LOO-DINKINS

Bachelor of Science in Forestry
University of New Brunswick
Fredericton, N.B., Canada
1979

Master of Science
Oklahoma State University
Stillwater, Oklahoma
1983

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 1986

COMPARISONS AMONG COMBINATIONS OF PROGENY

TEST DESIGNS AND SELECTION SYSTEMS FOR

MAXIMUM GENETIC GAIN IN LOBLOLLY

PINE (PINUS TAEDA L.)

Thesis Approved:

_____
Thesis Adviser

_____

_____

_____

_____
Dean of the Graduate College

ACKNOWLEDGEMENTS

I wish to extend my thanks to the Weyerhaeuser Company
in general, for funding my research and to Dr. Clements
Lambeth and Mr. Steve Duke, in particular. To Clem, I owe
thanks for sharing his ideas and for his continual
encouragement throughout the duration of the study and to
Steve, thanks for sharing his computer expertise.

Sincere gratitude is expressed to the members of my
committee; to Dr. Chuck Tauer, my major adviser, for his
ceaseless interest and numerous contributions, to Dr. Ron
McNew for his very useful input, particularly on the last
manuscript, to Dr. Laval M. Verhalen and to Dr. Richard R.
Frahm for their interest, encouragement and helpful advice
in preparation of the manuscripts.

Thanks are also due Dr. Stanley Carpenter, Head,
Department of Forestry, for finding ways to fill the
funding gap at crucial moments during the course of the
study.

Finally, and most of all, a loving thank-you to Randy,
my best friend and husband, for his encouragement, patience,
and occasional prodding, each one when it was most needed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Commercial forest land in southeastern Oklahoma and
southwestern Arkansas is often highly heterogeneous.
Progeny tests are designed to evaluate the performance of
offspring of selected trees to aid in the choice of parents
for the next generation of commercial seedlings. If the
progeny tests are to provide a realistic estimate of the
performance to be expected of commercial seedlings, they
must be established on a variety of commercial site types.
Thus, progeny test sites often contain a high degree of
variability.

Progeny tests usually serve the dual function of
providing information about the parent trees as well as
sources from which the next generation of parents are
selected. The selection process typically involves both
family mean and individual tree performance. On a variable
site, both the progeny test design and the method of
within-family selection may be very important.

Several approaches have been pursued in this study to
compare the statistical efficiencies of various designs, to
compare the relative genetic gains obtained by using
combinations of several designs and selection methods and
to quantify the environmental relationships among the
positions of trees in a plantation.

Three separate articles constitute Parts I, II and III. The first, "Statistical Efficiency of Six Progeny Test Designs on Three Loblolly Pine (<u>Pinus</u> <u>taeda</u> L.) Site Types", will be submitted for publication in the <u>Canadian Journal of Forest Research</u>. The second article, "Computer Simulations to Compare Combinations of Progeny Test Designs and Selection Systems for Maximum Genetic Gain in Loblolly Pine (<u>Pinus</u> <u>taeda</u> L.)", will be presented at a meeting of the International Union of Forest Research Organizations in Williamsburg, VA, in October, 1986. It will subsequently be published in the IUFRO proceedings. The third article, "Neighborhood Comparison as a Basis for Selection in Loblolly Pine (<u>Pinus</u> <u>taeda</u> L.)", will be submitted for publication in <u>Theoretical and Applied Genetics</u>.

PART I

STATISTICAL EFFICIENCY OF SIX PROGENY

TEST DESIGNS ON THREE LOBLOLLY PINE

(PINUS TAEDA L.) SITE TYPES

ABSTRACT

The environmental variances associated with family mean estimates for height were compared among six progeny test designs on three loblolly pine (_Pinus taeda_ L.) progeny test sites in southeastern Oklahoma and southwestern Arkansas. Twenty-five- (or 30-) tree block plots, 10-tree row plots, 5-tree noncontiguous plots, 2-tree row plots and single-tree plots were compared. The 5-tree noncontiguous, 2-tree and single-tree plot designs consistently resulted in lower environmental variances associated with family means than the larger plots. The noncontiguous plot design resulted in environmental variances as low as those for the single-tree plot design, except on the most variable site. The environmental variance associated with family means was as small with 30 trees per family arranged in single-tree or 5-tree noncontiguous plots as with 60 trees per family arranged in 5-tree row plots.

INTRODUCTION

Commercial loblolly pine (_Pinus taeda_ L.) sites in southeastern Oklahoma and southwestern Arkansas are usually highly variable. Progeny tests, designed to evaluate the breeding value of selected parent trees, must be established on a range of commercial site types. Thus, progeny tests in this region often contain a high degree of heterogeneity. Variability in a test site may appear as random microsite differences or as patches, that may be due to soil moisture, texture, or depth. Another common site type includes the variability associated with a gradient caused by distance from a water source, by a slope, or other factors. The precision in estimating family means in progeny tests depends upon the test design, and the importance of design efficiency may vary with site type.

Barnes and Schweppenhauser (1979) expressed the variance of a family mean determined from a replicated progeny test as:

$$Var(\bar{F}) = \sigma_f^2 + \sigma_p^2/r + \sigma_w^2/qr$$

where F = family mean, σ_f^2 = family variance component, σ_p^2 = among-plot variance component, σ_w^2 = within-plot variance component, r = number of replicates, and q = number of trees per plot. The part of the variance of a family mean that is largely attributable to environment is $\sigma_p^2/r + \sigma_w^2/qr$.

Thus, the environmental portion of the among-plot and within-plot variance can be substantially altered by changing the relative plot size and number of replicates.

The theoretical value of the replicated single-tree plot design is well known (Libby and Cockerham 1980, Wright 1973). Single-tree plots maximize the number of microsites sampled, and replication size is minimized. Possible confounding caused by environmental covariances among members of a family is eliminated. A 2-tree plot design should be almost as efficient as the single-tree plot design. The number of replicates would be halved, but replicate size would still be small. A noncontiguous plot design, with each "plot" consisting of a number of trees of each family randomly located throughout a replicate, should have most of the advantages of the single-tree plot design (Lambeth et al. 1983). In addition, the probability of any family being absent in a replicate is much lower.

A 4- to 6-tree row plot design has probably been used more frequently than the other designs in past and present progeny tests. Ten-tree row plot designs and designs with longer rows or block plots have also been used extensively in the past. Barnes and Schweppenhauser (1979) indicated that 10-tree row plots were used in Pinus patula Shiede et Deppe progeny tests established between 1967 and 1972 because the authors were concerned about possible non-normality of the genetic component of among-plot effects. They subsequently examined the theoretical relationships

between within- and among-plot variances and concluded that
a considerable increase in precision may result from using
3- to 5-tree row plots with more replicates. Evans et al.
(1961) used 25-tree plots because they noted that the
coefficient of variation became stable at approximately 20
trees. They were also concerned about possible non-
normality of within-plot effects with small plots. Franklin
(1971) noted, however, that the environmental effect is
always sufficiently large relative to the genetic values to
validate the assumption of normality. In addition the
genetic values themselves are usually normally distributed
(Falconer 1981).

The objective of this study was to compare the
statistical efficiencies in estimating family means for
height using 25- (or 30-) tree block plots, 10-tree row
plots, 5-tree row plots, 5-tree noncontiguous plots, 2-tree
plots and single-tree plots. The designs were compared on
three distinctly different site types in southeastern
Oklahoma and southwestern Arkansas.

MATERIALS AND METHODS

Height data from three Weyerhaeuser Company loblolly
pine progeny tests were used to estimate environmental
values for each site. The Dierks plantation, AOGC7505, is a
half-sib progeny test planted in 1975, located near Dierks
in southwestern Arkansas. The site is relatively uniform
with a slight incline. The test area is rectangular with 12
replicates, each consisting of a 15 X 18 tree block with 54
families. Eight of the replicates consist of 5-tree row
plots and the remaining four replicates consist of 5-tree
noncontiguous plots. The Cloudy plantation, AOWF8088, is a
full-sib progeny test planted in 1980, located near Cloudy
in southeastern Oklahoma. The test straddles a ridge
oriented in an east-west direction and six replicates are
located on each side of the ridge on a steep slope. The
replicates consist of 34 families of 5-tree row plots.
Filler plots were used on microsites that were judged to be
anomalous. The test shape is highly irregular. The
Hempwallace plantation, AOWF8085, is a full-sib progeny test
planted in 1980, located near Hempwallace in southwestern
Arkansas. The site is relatively wet, causing a patchy
pattern of variability. The test consists of 40 families
arranged in 12 replicates, eight of which have a 5-tree row
plot design and the remaining four have a 5-tree

noncontiguous plot arrangement. Filler trees were planted extensively on anomalous microsites, and the shape of the test is highly irregular. Eight-year data from the Dierks plantation and 5-year data from the other two plantations were used. The Dierks site was considered to be the most uniform; the Hempwallace site was judged to be intermediate and the Cloudy site was considered to be the most variable.

Residual values for height were obtained by subtracting the plantation mean and family effect from the height value for each tree in each plantation using the GLM procedure of the Statistical Analysis System (1982a). The residual values consisted of the remaining individual genetic effect and an effect associated with the microsite. Removal of full-sib family effects removes one-half of the additive genetic effect. In the case of half-sib families, one-quarter of the additive genetic effect is removed. It would be expected, however, that the environmental effect would be sufficiently larger than the remaining genetic effect to allow the residual values to primarily reflect the quality of the microsite, as suggested by Franklin (1971). If individual tree heritability is equal to .20, approximately 90% of the variance of the residual values in a full-sib progeny test would be due to nonadditive genetic and environmental causes.

The residual values became the environmental map for each plantation. Each position was given a row and column number and filler trees were indicated. As filler trees

were not measured in the plantations, they appeared as missing data in the mapped residual data sets. Values were approximated for filler trees using a random function of the Statistical Analysis System (1982b) to avoid the inflated apparent mortality. For each filler tree, a value was chosen at random from a normal distribution using the mean and variance of the surrounding 24 trees. Positions empty due to mortality were treated as missing data.

In comparing the six designs, it was important that each design fit in the same area of each plantation in such a way that replicates paralleled gradients to ensure valid comparisons. Because of these constraints, the number of usable positions varied among plantations. The rectangular Dierks plantation provided 2880 usable positions. The Hempwallace plantation had 1900 usable positions and only 1500 positions could be used from the Cloudy plantation. Each of the six designs was superimposed on the map of residual values using a computer model (Loo-Dinkins, Part II). Two variations in the number of families and number of trees within families were tested. Because the number of usable positions within each plantation was constant, increasing the number of families resulted in a proportionate decrease in the number of trees per family. The first set of conditions included 48 families with 60 trees per family on the Dierks plantation, 38 families with 50 trees per family on the Hempwallace plantation, and 30 families with 50 trees per family on the Cloudy plantation.

The numbers of families were increased to 96, 63, and 50 for Dierks, Hempwallace, and Cloudy, respectively, and each included 30 trees per family under the second set of conditions. The block plot design was not tested with the lower number of trees per family.

An analysis of variance was performed on the residual values with each of the superimposed designs to examine the partitioning of the variance components. The model used was:

$$\text{Residual} = \text{Rep} + \text{Plot(Rep)} + \text{Within Plot.}$$

The part of the total environmental variance that would contribute to the variance of a family mean, $\sigma_p^2/r + \sigma_w^2/qr$, denoted as E_F, was estimated for each site. The within and among plot variances would be subject to change among noncontiguous plot designs on the same environment depending upon the randomization pattern. Thus, for this design, the average values for σ_p^2 and σ_w^2 were estimated from 10 randomly assigned arrangements of noncontiguous plots for each site. Comparisons were made among the designs relative to the performance of the 5-tree row plot design, because this design is more frequently used than the others.

RESULTS

The environmental portion of the variance of a family mean, E_F, was larger for the large plot designs than for the designs with smaller plots (Tables 1, 2, and 3), as expected. Increasing the number of families to double or almost double those under the first set of conditions generally doubled or almost doubled E_F. The increase in the variance as the number of families was increased was greater in the case of 5-tree and 10-tree row plots than for the 5-tree noncontiguous, 2-tree plot, or single-tree plot designs, however.

The 25- (or 30-) tree block plots produced a higher E_F value than did any other design on each site. On the least variable site, the Dierks plantation, E_F was approximately three times larger for the 30-tree block plot design than for the 5-tree row plot design. E_F was five and 10 times greater for the 25-tree block plot design than for the 5-tree row plot design on the second most uniform site, the Hempwallace plantation, and on the most variable site , the Cloudy plantation, respectively. The block plot designs also performed substantially worse than the 10-tree row plot design on each site. The 10-tree row plot design consistently had the second highest E_F value when the number of families was at the lower number. The 10-tree row plot

design consistently had the highest E_F value when the higher number of families was tested, as block plots were only tested with the lower number of families. E_F values resulting from 10-tree row plot designs were consistently approximately 1.4 times greater than the E_F values for the 5-tree row plot design, regardless of site and number of families tested.

The 5-tree row plot design consistently had smaller E_F than block plot or 10-tree row plot designs, but larger E_F values than the 5-tree noncontiguous, 2-tree, or single-tree plot designs. The superimposition of the 5-tree row plot design on each of the plantations resulted in plots being placed on the same five positions as the original five-tree row plots. Thus, error in estimating family effects would result in higher among plot variance and lower within plot variance for the five-tree row plot design than could be attributed to the environment. To test this, mean correlation coefficients were estimated between residual values for each position within a moving five by five-tree area (Loo-Dinkins, Part III). Correlations between residual values within the original five-tree plot areas were no larger, on the average, than correlations between residuals from parallel plots. Thus, possible confounding effects in E_F estimates for five-tree plots are minimal.

The 5-tree row plot E_F ratio for low to high number of families (high to low number of tree per families) was greatest on the most variable site. On the most uniform

site, the Dierks plantation, the numbers of trees per family tested were 60 and 30 (2:1) and the E$_F$ ratio was 1:1.997. The E$_F$ ratio increased to 1:2.022 on the second most uniform site, the Hempwallace plantation, while the ratio of high to low number of trees per family decreased to 5:3. On the most variable site, the Cloudy plantation, the ratio of high to low number of trees per family tested was also 5:3, but the E$_F$ ratio increased to 1:2.184. The opposite trend was apparent for the five-tree noncontiguous plots, with ratios of E$_F$ values from high to low numbers of trees per family of 1:1.970, 1:1.714, and 1:1.667 for the most uniform to the most variable sites, respectively. Thus, relative to the 5-tree row plot design, the 5-tree noncontiguous plot design improved with increased numbers of trees per family on the sites with greatest variability.

The 5-tree noncontiguous plot design resulted in E$_F$ values that were always as small as or smaller than those for 5-tree row plots. The noncontiguous plot design also compared favorably with the 2-tree plot design, but except for the Hempwallace site with the low number of families, the single-tree plot design resulted in lower E$_F$ values. The difference was slight on the most uniform site (Table 1).

The 2-tree plot design always performed better than the 5-tree row plot design on each site regardless of number of families tested, but generally slightly worse than the 5-tree noncontiguous plot design. The 2-tree plot design did

not improve as much relative to the 5-tree row plot design when the number of families was increased as did the noncontiguous plot and single-tree plot designs.

The single-tree plot design generally yielded the lowest EF values in spite of missing plots due to a 5 to 15% mortality rate in the plantations which resulted in the absence of some families from some replicates. The among replicate difference was greatest when the higher number of families was tested; thus more variance was removed by the replicate effect for the higher number of families than for the lower number with single-tree plots. The single-tree plot EF values for the high number of families were similar or a little higher than those for the 5-tree row plot design with low number of families.

DISCUSSION

Statistical theory predicts that either increasing plot size or reducing replicate number increases the variance of a family mean (Cochran and Cox 1957). Also the site type would be expected to influence the magnitude of differences resulting from the various designs. Results of this study indicate that the site type may have a strong impact on the relative efficiencies of various plot sizes and replicate number.

Large block plots are sometimes used to estimate the yield performance of a family by harvesting the entire block. This design should only be used on uniform sites if a realistic estimate of family performance is to be obtained. The performance of the block plot design decreased drastically relative to that of the 5-tree row plot design on the more variable sites.

The difference in E_F values between the 5-tree and 10-tree row plot designs (approximately 1:1.4) did not change substantially with site or when the number of families was increased at the expense of the number of trees per family. In most cases, the magnitude of the individual variance components, σ_r^2, σ_p^2, and σ_w^2 did not greatly change between the 5-tree and 10-tree row plot designs. The difference in E_F values between the two designs appeared to result

16

primarily from the reduction by half of the divisor for σ_p^2 in E$_F$ values for the 10-tree row plot design. Each of the other designs responded differently on each site relative to the 5-tree row plot design, implying that other factors were more important than the magnitude of the divisor in those cases. On the most uniform site, more variance was removed by the replication with 10-tree row plots than with 5-tree row plots. On the more variable sites, this tended to be reversed. In general, the replicates containing 2- and single-tree plots removed more variance than in the case of larger plot designs, but replication appears to be much less important on the two more uniform sites than on the most variable site.

Five-tree row plots are extensively used in many tree improvement programs. They are relatively easy to lay out, map, and thin by family (Lambeth et al. 1983). They are also small enough to avoid the more serious problems of the larger plots. Results of this study indicate that the convenience of 5-tree row plot designs may not outweigh the increased efficiency of the 5-tree noncontiguous, 2-tree, or single-tree plot designs. Lambeth et al. (1983) reported that 60 individuals arranged in 5-tree noncontiguous plots would provide an estimate of a family mean as efficiently as 85 individuals arranged in 5-tree row plots. The ratio appears to be site specific. On the most uniform site, 30 trees per family arranged in 5-tree noncontiguous plots produced an E$_F$ value very similar to that of 60 trees per

family arranged in 5-tree row plots. This implies that, on a uniform site, 30 trees per family arranged in a 5-tree noncontiguous plot design may provide as accurate an estimate of family means as twice that number of trees per family in 5-tree row plots. On the less uniform sites, the noncontiguous plot designs with 30 trees per family performed at a level between 5- and 10-tree row plots with 50 trees per family. The difference may be explained by the observation that the ratio of within- to among-plot variance for the 5-tree row plot was much greater on the more variable sites than on the most uniform site. The noncontiguous plot design performed very well compared with the 5-tree row plot design on the most uniform site because the among-plot variance is much lower for the noncontiguous plot design. The ratio of within- to among-plot variance was always very high for noncontiguous plots. Because the divisor for within-plot variance is five times the divisor for among-plot variance in the expression for E_F, the within- to among-plot variance ratio is the primary difference between the efficiencies of row and noncontiguous plot designs. As the ratio of within- to among- plot variance increased from 3:1 to 41:1 for the 5-tree row plots from the most uniform site to the most variable site, the difference between the 5-tree row plot and 5-tree noncontiguous plot designs decreased.

The noncontiguous plot design generally produced lower E_F values than the 2-tree plot design except on the most

variable site. The EF values for noncontiguous and single-tree plot designs were similar on the two most uniform sites but the single-tree plot design was considerably better on the most variable site. The primary difference between the noncontiguous and single-tree plot EF values is in the amount of variance removed by the replication. On the two most uniform sites, the relative sizes of replicates were not reflected in the size of σ_r^2. Replication was much more important on the most variable site, as would be expected.

The single-tree plots generally out-performed the other designs, although on the most uniform sites, the noncontiguous plot design was a close second. The difference between single-tree plots and 5-tree row plots was considerably smaller on the most variable site than on the other two sites due to orientation of the replicates. The replicates paralleled the gradient and the family plots extended across the gradient. Thus, most of the within-replicate variance was contained within plots rather than among plots. The 5-tree row within-plot variance component would have the same divisor as the single-tree plot among-plot variance. Thus, the smaller the among- to within- plot variance ratio, the closer 5-tree row plots will be in performance to single-tree plots, regardless of the size of the replicate variance component.

CONCLUSIONS

A 5-tree noncontiguous, 2-tree, or single-tree plot design should always produce more accurate estimates of family means than 5-tree row plots. The difference between the 5-tree row plot design and the more efficient designs is greatest on sites with a random pattern of variability where a significant part of the within-replicate variability is due to-plot-to plot variability. On a site with a strong gradient, if the replication is done efficiently, most of the variance will be contained within family plots, allowing the 5-tree row plot design to compare more favorably with the more efficient designs. Accuracy of family mean estimates, on a relatively uniform site, may be as high with only 30 trees per family arranged in noncontiguous or single-tree plots as with 60 trees per family arranged in 5-tree row plots. This would allow twice as many families to be tested within the same area wihout loss of precision, allowing greater genetic gain per unit cost in progeny testing.

LITERATURE CITED

Barnes, R. D., and M. A. Schweppenhauser. 1979. Genetic
    control of 1,5-year-old traits in *Pinus patula* Schiede
    et Deppe and a comparison of progeny test methods.
    Silvae Genet. 28: 156-167.

Cochran, W. G., and G. M. Cox. 1957. Experimental Designs.
    John Wiley and Sons, Inc., New York, NY.

Evans, T. C., J. C. Barber, and A. E. Squillace. 1961. Some
    statistical aspects of progeny testing. Proc. Sixth
    South. Conf. For. Tree Impr. pp. 73-79.

Falconer, D. S. 1981. Introduction to quantitative genetics.
    2nd ed. Longman Inc. New York, NY.

Franklin, E. C. 1971. Statistical validity of single-tree
    plot in forest genetics research. Silvae Genet. 20: 73-
    75.

Lambeth, C. C., W. T. Gladstone and R. W. Stonecypher. 1983.
    Statistical efficiency of row and noncontiguous family
    plots in genetic tests of loblolly pine. Silvae Genet.

32: 24-28.

Libby, W. J. and C. C. Cockerham. 1980. Random non-contiguous

plots in interlocking field layouts. Silvae Genet.

29: 183-190.

Loo-Dinkins, J. A. Comparisons among combinations of progeny

test designs and selection systems for maximum genetic

gain in loblolly pine (Pinus taeda L) Part II. Unpubl.

Ph.D. dissertation. OSU, Stillwater, OK.

Loo-Dinkins, J. A. Comparisons among combinations of progeny

test designs and selection systems for maximum genetic

gain in loblolly pine (Pinus taeda L) Part III. Unpubl.

Ph.D. dissertation. OSU, Stillwater, OK.

Statistical Analysis System. 1982a. SAS user's guide :

Statistics. SAS Institute Inc. Cary, NC.

Statistical Analysis System. 1982b. SAS user's guide :

Basics. SAS Institute Inc. Cary, NC.

Wright, J. W. 1973. Genotype-environment interaction in north

central United States. For. Sci. 19: 113-123.

Table 1. Partitioning of the environmental portion of the variance of a family mean for height in loblolly pine, E_F, for the most uniform site, the Dierks plantation

Design	No. reps	σ_r^2	σ_p^2	σ_w^2	E_F	[a]Rel value
48 family plots/rep						
30-tree blocks	2	.1200	.1283	.8333	.0887	2.99
10-tree rows	6	.1312	.1690	.7410	.0405	1.36
5-tree rows	12	.1247	.2191	.6858	.0297	1.00
5-tree nc plots	12	.1284	.0050	.9009	.0154	0.52
2-tree plots	30	.1378	.1326	.7527	.0166	0.56
Single-tree plots	60	.1596	.8619		.0144	0.48
96 family plots/rep						
10-tree rows	3	.1589	.1725	.7410	.0822	1.39
5-tree rows	6	.1378	.2185	.6858	.0593	1.00
5-tree nc plots	6	.1398	.0000	.9135	.0304	0.51
2-tree plots	15	.1734	.2237	.6291	.0359	0.61
Single-tree plots	30	.1844	.8405		.0280	0.47

[a]$E_F = \sigma_p^2/r + \sigma_w^2/rq$, where σ_p^2 is the among plot variance, r is the number of replicates, σ_w^2 is the within plot variance, and q is the number of individuals in a plot.
[b]σ_r^2 is the among replicate variance.
[c]Relative value refers to the size of E_F for each design relative to E_F for the 5-tree row plot design.

Table 2. Partitioning of the environmental portion of the variance of a family mean for height in loblolly pine, $^aE_{\bar{F}}$, for the second most uniform site, the Hempwallace plantation.

Design	No. reps	$^b\sigma_r^2$	σ_p^2	σ_w^2	$E_{\bar{F}}$	^{c}Rel value
38 family plots/rep						
25-tree blocks	2	.0692	.0415	.1533	.0238	5.12
10-tree rows	5	.0643	.0176	.1600	.0067	1.44
5-tree rows	10	.0655	.0156	.1544	.0046	1.00
5-tree nc plots	10	.0659	-.0034	.1730	.0035	0.74
2-tree plots	25	.0648	.1160	.1641	.0037	0.81
Single-tree plots	50	.0627	.1675		.0036	0.78
63 family plots/rep						
10-tree rows	3	.0652	.0247	.1602	.0135	1.45
5-tree rows	6	.0581	.0250	.1548	.0093	1.00
5-tree nc plots	6	.0584	.0003	.1792	.0060	0.64
2-tree plots	15	.0651	.0292	.1601	.0066	0.70
Single-tree plots	30	.0641	.1662		.0055	0.59

$^aE_{\bar{F}} = \sigma_p^2/r + \sigma_w^2/rq$, where σ_p^2 is the among plot variance, r is the number of replicates, σ_w^2 is the within plot variance, and q is the number of individuals in a plot. $^b\sigma_r^2$ is the among replicate variance.
^{c}Relative value refers to the size of $E_{\bar{F}}$ for each design relative to $E_{\bar{F}}$ for the five-tree row plot design.

Table 3. Partitioning of the environmental portion of the variance of a family meanfor height in loblolly pine, EF, for the most variable site, the Cloudy plantation.

Design	No. reps	σ_r^2	σ_p^2	σ_w^2	EF	aRel. value
30 family plots/rep						
30-tree blocks	2	.0408	.0916	.2071	.0499	10.21
10-tree rows	5	.0814	.0122	.2403	.0072	1.47
5-tree rows	10	.1045	.0053	.2183	.0049	1.00
5-tree nc plots	10	.1045	.0044	.2191	.0048	0.98
2-tree plots	25	.1551	.0238	.1449	.0038	0.79
Single-tree plots	50	.1386	.1821		.0036	0.74
50 family plots/rep						
10-tree rows	3	.0852	.0207	.2403	.0149	1.39
5-tree rows	6	.0944	.0207	.2183	.0107	1.00
5-tree nc plots	6	.0948	-.0005	.2413	.0080	0.75
2-tree plots	15	.1267	.0433	.1563	.0081	0.75
Single-tree plots	30	.1524	.1703		.0057	0.53

aEF = $\sigma_p^2/r + \sigma_w^2/rq$ where σ_p^2 is the among plot variance, r is the number of replicates, σ_w^2 is the within plot variance, and q is the number of individuals in a plot.
bσ_r^2 is the among replicate variance.
cRelative value refers to the size of EF for each design relative to EF for the five-tree row plot design.

PART II

COMPUTER SIMULATIONS TO COMPARE COMBINATIONS

OF PROGENY TEST DESIGNS AND SELECTION

SYSTEMS FOR MAXIMUM GENETIC GAIN

IN LOBLOLLY PINE (PINUS

TAEDA L.)

ABSTRACT

Six progeny test designs in combination with four within family selection systems were compared on three loblolly pine (_Pinus_ _taeda_ L.) progeny test sites in southeastern Oklahoma and southwestern Arkansas. Large block plots, 5- and 10-tree row plots, 5-tree noncontiguous plots, 2-tree plots, and single-tree plots were tested in combination with selection based on deviations from replicate means, deviations from neighborhood means, a combination of replicate and neighborhood deviations, and deviations from plot means. Large block plots performed consistently poorly compared with the other designs, and selection based on deviations from a plot mean was always poor compared with the other selection systems. The single-tree plot design in combination with replicate or neighborhood deviation selection performed best on all sites. The 5-tree noncontiguous and 2-tree plot designs generally performed better than the larger row or block plots, particularly when the number of families tested was increased at the expense of number of trees per plot.

INTRODUCTION

Loblolly pine (<u>Pinus taeda</u> L.) is the most important commercial forest species in southeastern Oklahoma and southern Arkansas. The species is often planted on sites that are highly variable. This variability may be due to such factors as slope, aspect, soil moisture conditions, soil texture and soil fertility. Commercial forest sites typically include much more heterogeneity than do agricultural fields. The heterogeneity may appear as random microsite variability; it may be characterized by larger patches that grade into one another; or it may be in the form of a definite gradient.

The primary purpose of progeny testing is to evaluate family performance in order to rank potential parents of the next generation of commercial seedlings. Progeny tests must sample a range of commercial planting sites in order to predict the commercial performance of a family. As the size and purpose of progeny tests does not allow avoidance of much of the variability, the test sites may be highly variable. A secondary but equally important function of many progeny tests is to provide a base for selection of the next generation of parents. Thus, in addition to accurate estimation of family means, which depends primarily on the test design, an optimal within-family

selection system is important.

PROGENY TEST DESIGNS

The variance associated with a family mean estimate is largely determined by among-plot and within-plot variance (Barnes and Schweppenhauser 1979). The among-plot variance can be substantially altered by changing the relative plot size and number of replicates. Given a fixed plantation size and number of families tested, size of plots will be inversely related to the number of replicates. In addition, the number of trees per family will vary inversely with the number of families that can be tested. On a variable site, the relative size and arrangement of plots largely determines the statistical efficiency of the test (Libby and Cockerham 1980, Loo-Dinkins, Part I).

Statistical theory indicates that the smaller the plot size, the more efficient it should be (Libby and Cockerham 1980). Single tree plots maximize the number of microsites sampled and minimize replicate size (Wright 1973). Possible confounding of environmental and genetic covariances among members of a genetic unit is eliminated. In practice, use of single tree plots may lead to many missing cells due to the combined effects of thinning and unplanned mortality. A 2-tree plot design would also ensure thorough microsite sampling by each family and replicates would be small, although twice as large as replicates of single-tree plots.

A noncontiguous plot design, with each plot consisting

of a number of trees randomly located throughout a
replicate, would have most of the advantages of the single-
tree plot design. In addition, the probability of a given
family being absent from one or more replicates is much
lower (Lambeth et al. 1983, Libby and Cockerham 1980).
Lambeth et al. (1983) compared a test design consisting of
5-tree noncontiguous plots with a 5-tree row plot design
and found that among plot variance was greater for row
plots than for noncontiguous plots. Thus the number of
replicates required to obtain a given level of precision is
lower for noncontiguous plots than for row plots.

Four- to 6-tree row plots constitute the most
extensively used design in todays forest tree improvement
programs. Field layout and silvicultural thinning by
family are relatively easy. Plots are large enough that
analysis on a plot means basis usually would not result in
missing cells. Larger row plots and block plots have been
used frequently in the past in progeny test designs. Early
tree-improvement workers were concerned about possible
non-normality of among- or within-plot effects with small
plot size (Evans et al. 1961, Barnes and Schweppenhauser
1980). Franklin (1971) noted that the environmental effect
is always sufficiently large relative to the genetic values
to validate the assumption of normality.

Loo-Dinkins (Part I) compared the theoretical
efficiencies of a number of progeny test designs on three
site types. She found that the relative efficiency of

designs varied with the site type, but that single-tree, two-tree and noncontiguous plot designs always resulted in lower variance associated with family means than larger row or block plots. Her results indicated that, on a relatively uniform site, 30 trees per family, arranged in noncontiguous or single-tree plots, may provide as accurate an estimate of family means as would 60 trees per family arranged in 5-tree row plots.

WITHIN FAMILY SELECTION

Expressions for expected genetic gain have been derived and are well known for the traditional selection systems (Shellbourne 1969, Namkoong 1979). Some combination of among-and within-family selection is usually utilized when selecting in a progeny test. The optimal combination of among- and within-family selection depends primarily upon the heritability of the trait under improvement (Falconer 1981). In practice, the relative degree of within- and among-family selection is also influenced by inbreeding considerations. Frequently, the best families are identified and one or two individuals within each family are selected. The optimal method of selection of individuals within families may vary with design and site type.

Current within-family selection methods include use of deviations from plot means and deviations from replicate means in the initial stages, to identify trees worthy of a field assessment. The Western Gulf Forest Tree Improvement

Cooperative selects trees from progeny tests using deviations from plot means among other considerations. The North Carolina State Forest Tree Improvement Cooperative includes deviations from a replicate mean as part of the selection criteria within progeny tests.

The optimal size and shape of a neighborhood comparison area may fall between a plot and a replicate. Gardner (1961) described a method to reduce the environmental variance within the selection area for corn. He superimposed a grid on the test area and selected a set number of individuals from within each square. A similar procedure that would be applicable to selection of forest trees in a progeny test would be to superimpose a grid on the plantation to create neighborhoods. Within-family selections would then be made on the basis of deviations from neighborhood means. The neighborhoods should be small enough that on the average the environmetal correlations between positions would be greater than zero. The size and shape of neighborhoods would vary with site type (Loo-Dinkins, Part III).

The theory of progeny test design efficiency and expected gains for various selection systems have been examined, but not in combination. Combinations of progeny test designs and within-family selection systems can be compared on various site types, using a simulation model. A simulation model imposes limitations on the kinds of selection systems that can be tested; and results provide

guidelines, only as good as the assumptions used in building the model. Modeling may be the only way to compare different designs on one site, however, and the technique may provide new insights.

The objective of the study was to compare the influence of various progeny test designs in combination with several within-family selection methods on the genetic gain in 5- or 8-year height in loblolly pine. The efficiency of the various combinations was compared using an environmental data base generated from three progeny test sites with superimposed genetic populations generated by a simulation model.

MATERIALS AND METHODS

Environmental data were obtained from three
Weyerhaeuser Company loblolly pine progeny tests occupying
distinctly different site types in southeastern Oklahoma
and southwestern Arkansas. The shape and layout of
replicates for each plantation is shown in Fig. 1. Dierks
plantation, AOGC7505, is a half-sib progeny test, planted
in 1975, located near Dierks, Arkansas. The site is
relatively uniform with apparently random microsite
variability. The test area is rectangular with 12
replicates each consisting of a 15 X 18 tree block with 54
families. Eight of the replicates consist of 5-tree row
plots. The remaining four replicates contain 5-tree
noncontiguous plots. The Cloudy plantation, AOWF8088, is a
full-sib progeny test planted in 1980, located near Cloudy,
Oklahoma. The test contains a strong gradient as it
straddles a ridge with an east-west orientation. Six
replicates consisting of 34 5-tree family row plots are
located on each side of the ridge. Filler plots were used
on microsites which were judged to be anomalous. The test
shape is highly irregular (Fig. 1). The Hempwallace
plantation, AOWF8085, is a full-sib progeny test planted in
1980, located near Hempwallace, Arkansas. The site is
relatively wet, causing a pattern of patchy variability.

The test consists of 40 families arranged in 12 replicates;
eight of which have a 5-tree row plot arrangement. The
remaining four replicates have a 5-tree noncontiguous plot
arrangement. Filler trees were planted extensively on
anomalous microsites and the shape of the test is highly
irregular. Eight-year height data were used from the
Dierks plantation, and 5-year height data were used from
the other two plantations. The survival in the three
plantations ranged from 85 to 95%.

Residual values for height were obtained by
subtracting the family effect and overall plantation mean
from the height of each tree using the GLM procedure of the
Statistical Analysis System (1982a). The resultant
deviations with a mean of zero included a portion of the
within family genetic effects in addition to the true
environmental effects. The environmental effects were assumed
to be sufficiently larger than the remaining genetic effects
to allow the residual value to primarily reflect a measure
of the microsite value, as suggested by Franklin (1971).
For example, if individual tree heritability is .20,
approximately 90% of the variance of the residual values in
a full-sib progeny test would be due to nonadditive genetic
and environmental causes.

The environmental residual data sets were mapped using
a row and column coordinate system. The original 5-tree
plots in each plantation fell within columns. As filler
trees were not measured in plantations, they appeared as

missing values in the mapped residual data sets. Thus they were indistinguishable from trees missing because of mortality. To avoid the problem of an apparent inflated mortality rate, environmental values were approximated for the filler positions. The mean and variance of neighboring 24 trees, including two trees on each side of the filler position, were used to generate approximate environmental values for the filler positions with a random normal function of the Statistical Analysis System (1982b) A grid pattern was superimposed on each data set to establish 5 X 5-tree neighborhoods for neighborhood selection.

Six planting designs were tested on each environmental data set with a relatively low number of families (30 to 48) and a high number of trees per family (50 or 60). Block plots, 5 X 6 trees on the Dierks plantation, and 5 X 5 trees on the other two plantations were arranged in two replicates. The block plots could not be tested when the number of families was increased (50, 63, or 96) and number of trees per family was reduced to 30. The remaining five designs were tested under both sets of conditions. Ten-tree row plots, 5-tree row plots, 5-tree noncontiguous plots, 2-tree row plots, and single-tree plots were arranged in the appropriate number of replicates for each data set. The within-family selection systems included selection based on deviations from a replicate mean, deviations from a plot mean, deviations from a neighborhood mean, and a combination of replicate and neighborhood mean deviations.

The combination selection method involved choosing the best ten trees from each selected family on the basis of deviations from the replicate mean and the best two trees from within the ten were then chosen on the basis of neighborhood deviations.

A model was constructed using SAS Macro Language (1982b) to generate genetic populations with an among- and within-family variance structure to correspond to a heritability of .20 or .30, and to superimpose each genetic population, arranged in each of the designs on the environmental data sets. These heritabilities were chosen because they are within the range of reported heritability values for height in loblolly pine (Zobel and Talbert 1984). The genetic structure simulated half-sib families with 3/4 of the additive genetic variance occurring within families for the Dierks plantation. The genetic structure simulated full-sib families for the other two plantations. The original plantation mean was added to the additive genetic and environmental values for each design to result in simulated height values corresponding to each position. The ten best families were chosen based on ranking of family means and the within family selection systems were applied to choose two phenotypes from each of the chosen families.

The model was iterated 100 times for each set of conditions. The two heritability levels (.20 and .30) were modeled at a high and low number of families on each of the

three data sets. The test design-selection system combinations were compared by examining the mean simulated genetic value of the selected phenotypes relative to the maximum possible mean genetic value. The maximum possible genetic value was calculated as the mean genetic value of the actual two best genotypes from the 10 families with the highest family genetic values. The relative performance of a design-selection system combination was then expressed as the percentage of the maximum that was attained. The mean percentages were calculated over 100 iterations. The relative proportions of the genetic value of selected trees due to family versus individual tree genetic effects were determined under some sets of conditions.

RESULTS

The differences among progeny test design-selection system combinations, expressed as the mean percentage of the maximum genetic value that could be attained, are less than might have been expected based on relative statistical efficiencies of the designs (Loo-Dinkins, Part I, Lambeth et al. 1983) (Tables 1-12). The percentage of the maximum possible mean genetic value of the 20 selected phenotypes was consistently between 50 and 75%, regardless of test design, selection system, site type, or heritability. The standard errors associated with the mean percentage for 100 iterations were consistently small, ranging from 0.50 to 0.95. Thus, small differences, on the order of 1.6 to 2.5 percentage points, between percentages of the maximum possible mean genetic value could be declared statistically significant. Differences among designs, among selection systems across designs, and among design-selection system combinations were found to be highly significant using an analysis of variance. The differences between design-selection system combinations decreased when heritability was increased from .20 to .30. The percentage of the maximum genetic value increased by five to 10 points with increased heritabilities and the increase was greatest for designs with lowest percentages of maximum genetic value.

On the two most variable sites, Cloudy and Hempwallace, the percentage of the maximum genetic value increased by eight to 10 points with the 25-tree block plot design but only by five to six points for the single-tree plot design when heritability was increased from .20 to .30.

The large block plot designs consistently performed poorly in comparison with the other designs regardless of the selection system. Differences among the other designs were slight in most cases but a trend for small plots to perform better than larger plots was apparent under each set of conditions for each site. The 5-tree row plot design was almost always slightly better then the 10-tree row plot design and the noncontiguous, single-tree and 2-tree plot designs almost always performed better than the 5-tree row plot design. The trend was strongest when the lower number of trees per family with the higher number of families was tested.

Selection based on deviations from a plot mean consistently resulted in lower percentages of the maximum genetic value than any of the other selection systems. This method performed particularly poorly in combination with the 2-tree plot design. Selection based on deviations from plot means performed best relative to the other selection systems when the design consisted of large block or row plots. As plot size decreased the difference between percentages resulting from plot deviation selection and percentages resulting from the other selection methods

increased. Selection based on deviations from a replicate mean generally improved as plot size, and consequently, replicate size decreased. Selection based on deviations from a neighborhood mean was always better than plot mean deviation selection and generally slightly better than selection based on deviations from a replicate mean when the plot size was greater than five trees.

When the number of families was increased at the expense of the number of trees per family, differences between the designs were greater. Generally the row plot designs performed slightly worse and the 5-tree noncontiguous, single-tree and 2-tree plot designs resulted in slightly higher percentages of the maximum genetic value than were attained with the smaller number of families.

The average proportion of the genetic value of selected phenotypes that was due to the individual genetic value relative to the family genetic value varied from 0.40 to 0.60 (Tables 13 and 14). The proportions for the Dierks site were higher than for the other two plantations because the genetic structure in the Dierks plantation is half-sib, thus three-quarters of the genetic variance is within family. The other two plantations have a full-sib family structure, thus one-half of the genetic variance would be within family.

In general, the designs showing the highest proportions attributable to individual genetic values are those which attained lowest percentages of the maximum genetic value.

Within designs, however, the selection systems that
resulted in the highest individual proportions of the
genetic value also had highest percentages of the maximum
genetic value. The proportion of the mean genetic value of
selections that was due to the individual genetic value
increased by approximately 0.10 when the number of families
in the Cloudy plantation decreased from 50 to 30, and
number of trees per family was increased from 30 to 50
(Tables 13 and 14). The number of families chosen remained
constant at 10, allowing the selection intensity within
families to increase from 1/15 to 1/25, while the among
family selection intensity decreased from 1/5 to 1/3 when
the number of families was reduced. The overall intensity
did not change but the composition of the mean genetic
value shifted substantially in the individual direction.
This shift was accompanied by a decrease in the percent of
the maximum genetic value attained by the most efficient
designs. Increasing the heritability from .20 to .30
resulted in a slight increase in the proportion of the
genetic value that was due to individual values for the
Cloudy site. This was not examined on the other sites.

DISCUSSION

Statistical theory indicates that differences among
the designs, regardless of the selection system, could be
expected to be greater than those found in this study
(Lambeth et al. 1983, Loo-Dinkins, Part I). The designs
showed greater differences when number of families was
increased and number of trees within family was decreased.
It appears that given 50 or 60 trees per family and 30 to
50 families, the design is much less important, on the
sites tested, than when the number of trees per family is
reduced to 30. In particular, the difference between 5-
and 10-tree row plot designs is significant only when 30
trees per family were tested with a high number of
families.

The percentage of the maximum genetic value that was
attained with 5-tree noncontiguous, single, and 2-tree plot
designs generally increased when the number of families was
increased and number of trees per family decreased. The
percentage attained using 5- and 10-tree row plot designs
generally remained approximately the same or decreased
slightly when the number of families was increased and
number of trees per family decreased. This trend was
accompanied by a general decrease in the proportion of the
mean genetic value that reflected the individual value,

when the higher number of families was tested. Thus, at
the heritability values tested, the increased number of
families with fewer trees per family, allowing higher
among-family selection intensity, led to greater genetic
gains when the designs that allowed most accurate estimates
of family genetic values were used. The higher among
family selection intensity did not result in higher gains,
however, when the less efficient designs were used. This
result implies that if the more efficient designs are used,
a greater number of families may be tested in the same
space, resulting in increased genetic gain as well as
increased economic efficiency.

Selection based on deviations from a plot mean
consistently resulted in lower percentage of the maximum
genetic value than for each of the other selection systems.
As the number of trees per plot increased, the percentage
of the maximum genetic value attained by using plot
deviation selection increased to a plot size of 10 trees,
then decreased with 25- (or 30-) tree block plots (Fig.
2). The selection systems based on deviations from
replicate or neighborhood means yielded lower percentages
of the maximum genetic value as plot size increased.
Replicate deviation selection appeared slightly better with
single and 2-tree plots, and neighborhood deviation
selection performed better, on the average, for the block
plot designs. Five- and 10-tree row plot designs did not
differ for selection based on deviations from replicate and

neighborhood means when the relatively large number of trees per family was tested. Differences between the 5- and 10-tree row plot design percentages were apparent with the replicate and neighborhood deviation selection systems when 30 trees per family were tested (Fig. 2). There does not appear to be a difference between the 5- and 10-tree row plot designs, however, when the selection method was based upon deviations from plot means at the higher number of families. These results imply that with the larger number of families, the improved selection accuracy that could be expected of the plot deviation method when the plot size increased was offset by the negative influence of the larger replicate size. The size of the 10-tree row plot replicates approximately doubled when the number of families was increased, in each case, resulting in less efficient estimates of family value relative to the 5-tree row plot design.

The percent of the maximum genetic value, averaged over the three sites for the 5-tree noncontiguous design, was very close to the 2-tree plot results when the replicate and neighborhood deviation selection systems were used. The mean percentage value for noncontiguous plots was similar to that for 5-tree row plots with the relatively low number of families, and better than any of the other designs tested with the high number of families, when selection was based on deviations from plot means.

It was expected that the progeny test design-within

family selection system combination would be most important
on the most variable site. The Cloudy site appeared to
have the greatest amount of variability, with the
plantation slanting down-slope on either side of a ridge.
Differences between 10-tree row plot, 5-tree row plot and
5-tree noncontiguous plot designs were less than for the
other designs with 30 trees per family, however.
Selections at the Cloudy site had the lowest mean
proportion of individual vs. family genetic value (Table
13). Thus, accurate estimates of family means appeared to
be more important in determining the mean genetic value of
selected trees at the Cloudy site than at the other sites.
Replicates of 5- and 10-tree row plots on the Cloudy site
were arranged so that the family plots extended across the
gradient. As most of the variation in the site was due to
the gradient (Loo-Dinkins, Part III), each family row plot
sampled most of the variation within a particular
replicate, allowing accurate estimates of family genetic
value. Thus 5- or 10-tree row plot designs may be very
efficient on sites in which a strong, identifiable gradient
is the main source of variation, if the plots extend across
the gradient.

CONCLUSIONS

Differences in efficiency of estimation of family genetic value and selection accuracy among designs which include 10 trees or less in replicated family plots do not appear to be important if 50 or 60 trees per family are included. When the number of trees per family is reduced to 30 and the number of families tested is increased, the single-tree, 2-tree, and 5-tree noncontiguous plot designs are significantly better than the larger row plot designs. The single-tree plot design consistently performed the best under these conditions in spite of a 5-15% level of mortality resulting in some missing plots.

The optimum neighborhood size for selection may be greater than the 25-tree neighborhoods that were tested, but is smaller than the size of a replicate with 5-tree plots and 30 or more families. Selection based on deviations from a replicate mean is as good as selection based on deviations from a 25-tree neighborhood mean when replicates contained family plots with 10 trees or less.

Selection based on either deviations from a replicate mean or deviations from neighborhood means produces higher genetic gains than selection based on deviations from the plot mean, especially when plot size is five trees or less.

LITERATURE CITED

Barnes, R.D. and M.A. Schweppenhauser. 1979. Genetic
control of 1,5-year-old traits in *Pinus patula* Schiede
et Deppe and a comparison of progeny test methods.
Silvae Genet. 28:156-167.

Evans, T.C., J.C. Barber and A.E. Squillace. 1961. Some
statistical aspects of progeny testing. Proc. Sixth
South. Conf. Forest Tree Impr. pp. 73-79.

Falconer, D.S. 1981. Introduction to quantitative genetics.
2nd ed. Longman Inc. New York, NY.

Franklin, E.C. 1971. Statistical validity of single-tree
plot in forest genetics research. Silvae Genet. 20:73-
75.

Gardner, C.O. 1961. An evaluation of effects of mass
selection and seed irradiation with thermal neutrons on
yield of corn. Crop Sci. 1:241-245.

Lambeth, C.C., W.T. Gladstone, and R.W. Stonecypher. 1983.
Statistical efficiency of row and noncontiguous family

plots in genetic tests of loblolly pine. Silvae Genet. 32:24-28.

Libby, W.J., and C.C. Cockerham. 1980. Random non-contiguous plots in interlocking field layouts. Silvae Genet. 29:183-190.

Loo-Dinkins, J.A. Comparisons among combinations of progeny test designs and selection systems for maximum genetic gain in loblolly pine (Pinus taeda L.) Part I. Unpubl. Ph.D. dissertation. OSU, Stillwater, OK.

Loo-Dinkins, J.A. Comparisons among combinations of progeny test designs and selection systems for maximum genetic gain in loblolly pine (Pinus taeda L.) Part III. Unpubl. Ph.D. dissertation. OSU, Stillwater, OK.

Namkoong, G. 1979. Introduction to quantitative genetics in forestry. USDA For. Serv. Tech. Bull. No. 1588.

Shellbourne, C.J.A. 1969. Tree breeding methods. New Zealand For. Serv. Tech. Paper No. 55.

Statistical Analysis System. 1982a. SAS user's guide : Statistics. SAS Institute Inc. Cary, NC.

Statistical Analysis System. 1982b. SAS user's guide :

Basics. SAS Institute Inc. Cary, NC.

Wright, J.W. 1973. Genotype-environment interaction in north

central United States. For. Sci. 19:113-123.

Zobel, B., and J. Talbert. 1984. Applied forest tree

improvement. John Wiley and Sons, Inc., New York, NY.

Table 1. Mean percent of maximum genetic value for six designs and four selection systems on the Dierks plantation with 60 trees per family and h^2 = .2.

Design	Selection System			
	Rep Dev	Comb	N'hood	Plot
30-tree block plots	53.9	55.8	56.2	52.5
10-tree row plots	59.6	60.2	59.7	56.6
5-tree row plots	59.7	60.3	59.5	53.3
5-tree noncontiguous plot	59.6	61.3	60.6	53.9
2-tree plots	61.2	60.6	60.5	47.2
single tree plots	62.0	60.2	60.1	-

Standard errors ranged from 0.70 to 0.95.

Table 2. Mean percent of maximum genetic value for six designs and four selection systems on the Dierks plantation with 60 trees per family and h^2 = .3.

Design	Selection System			
	Rep dev	Comb	N'hood	Plot dev
30-tree block plots	62.0	65.3	64.1	61.0
10-tree row plots	65.9	66.3	65.6	62.6
5-tree row plots	66.5	66.6	66.3	59.9
5-tree noncontiguous plots	67.6	69.6	69.3	60.6
2-tree plots	68.3	67.7	67.3	53.3
single tree plots	69.8	68.7	68.5	-

Standard errors ranged from 0.50 to 0.83.

Table 3. Mean percent of maximum genetic value for five designs and four selection systems on the Dierks plantation with 30 trees per family and h^2 = .2.

Design	Selection System			
	Rep dev	Comb	N'hood	Plot dev
10-tree row plots	56.9	55.3	56.8	53.8
5-tree row plots	58.2	57.8	58.3	53.9
5-tree noncontiguous plots	61.5	60.8	61.8	57.5
2-tree plots	62.3	61.0	62.3	51.4
single tree plots	63.1	62.1	63.3	-

Standard errors ranged from 0.63 to 0.89.

Table 4. Mean percent of maximum genetic value for five designs and four selection systems on the Dierks plantation with 30 trees per family and h^2 = .3.

Design	Selection System			
	Rep dev	Comb	N'hood	Plot dev
10-tree row plots	67.0	66.0	65.6	64.1
5-tree row plots	68.0	69.5	69.5	63.7
5-tree noncontiguous plots	67.6	69.5	69.4	63.5
2-tree plots	70.0	70.7	70.6	60.6
single tree plots	69.4	70.5	70.9	-

Standard errors ranged from 0.68 to 0.79.

Table 5. Mean percent of maximum genetic value for six designs and four selection systems on the Cloudy plantation with 50 trees per family and h^2 = .2.

Design	Selection System			
	Rep dev	Comb	N'hood	Plot dev
25-tree blocks	55.4	55.1	55.4	54.8
10-tree row plots	63.4	62.0	63.1	58.3
5-tree row plots	62.7	63.6	62.7	57.6
5-tree noncontiguous plots	62.2	63.1	62.3	55.8
2-tree plots	64.6	66.1	64.3	54.2
Single-tree plots	65.0	65.9	64.5	-

Standard errors ranged from 0.60 to 0.94.

Table 6. Mean percent of maximum genetic value for six designs and four selection systems on the Cloudy plantation with 50 trees per family and h^2 = .3.

Design	Selection System			
	Rep dev	Comb	N'hood	Plot dev
25-tree block plots	65.4	64.3	64.9	63.3
10-tree row plots	69.5	67.8	69.4	63.4
5-tree row plots	69.7	70.6	69.6	62.8
5-tree noncontiguous plots	70.9	71.4	70.7	63.1
2-tree plots	69.8	72.0	69.5	59.9
single tree plots	71.2	71.8	70.5	-

Standard errors ranged from 0.50 to 0.78.

Table 7. Mean percent of maximum genetic value for five designs and four selection systems on the Cloudy plantation with 30 trees per family and h^2 = .2.

| | Selection System | | | |
Design	Rep dev	Comb	N'hood	Plot dev
10-tree row plots	65.5	65.6	65.7	62.8
5-tree row plots	65.6	65.5	65.4	61.3
5-tree noncontiguous plots	67.6	67.0	67.3	60.5
2-tree plots	67.5	67.4	66.6	58.1
single tree plots	70.0	68.5	68.3	-

Standard errors ranged from 0.61 to 0.79.

Table 8. Mean percent of maximum genetic value for five designs and four selection systems on the Cloudy plantation with 30 trees per family and h^2 = .3.

| | Selection System | | | |
Design	Rep dev	Comb	N'hood	Plot dev
10-tree row plots	70.8	71.1	71.1	67.2
5-tree row plots	72.7	71.7	71.7	66.8
5-tree noncontiguous plots	73.9	73.2	73.2	66.9
2-tree plots	73.2	73.4	73.2	63.1
single tree plots	74.7	73.1	73.0	-

Standard errors ranged from 0.51 to 0.64.

Table 9. Mean percent of maximum genetic value for six designs and four selection systems on the Hempwallace plantation with 50 trees per family and h^2 = .2.

| | Selection System | | | |
Design	Rep dev	Comb	N'hood	Plot dev
25-tree block plots	54.8	55.9	56.1	56.1
10-tree row plots	61.4	62.0	62.1	59.1
5-tree row plots	61.0	61.9	61.8	58.0
5-tree noncontiguous plots	62.5	63.1	63.2	58.0
2-tree plots	62.2	62.9	62.7	52.0
single tree plots	62.3	62.9	63.1	-

Standard errors ranged from 0..63 to 0.81.

Table 10. Mean percent of maximum genetic value for six
designs and four selection systems on the Hempwallace
plantations with 50 trees per family and h^2 = .3.

Design	Selection System			
	Rep dev	Comb	N'hood	Plot dev
25-tree block plots	63.4	64.6	64.7	64.6
10-tree row plots	67.5	67.9	68.1	64.7
5-tree row plots	66.7	67.3	67.1	62.7
5-tree noncontiguous plots	67.0	67.6	67.4	61.4
2-tree plots	66.6	66.6	66.6	56.5
single tree plots	67.6	68.0	68.0	-

Standard errors ranged from 0.54 to 0.80.

Table 11. Mean percent of maximum genetic value for five
designs and four selection systems on the Hempwallace
plantation with 30 trees per family and h^2 = .2.

Design	Selection System			
	Rep dev	Comb	N'hood	Plot dev
10-tree row plots	61.0	60.5	61.1	59.4
5-tree row plots	64.4	62.8	64.5	61.6
5-tree noncontiguous plots	64.9	64.5	64.9	61.1
2-tree plots	65.0	64.6	64.8	56.1
single tree plots	65.6	64.8	65.5	-

Standard errors ranged from 0.62 to 0.79.

Table 12. Mean percent of maximum genetic value for five
designs and four selection systems on the Hempwallace
plantation with 50 trees per family and h^2 = .3.

Design	Selection System			
	Rep dev	Comb	N'hood	Plot dev
10-tree row plots	67.9	69.1	69.1	67.9
5-tree row plots	69.8	70.6	70.9	67.2
5-tree noncontiguous plots	69.9	71.1	71.1	66.3
2-tree plots	73.1	72.9	72.8	62.9
single tree plots	71.5	72.1	72.1	-

Standard errors ranged from 0.53 to 0.67.

Table 13. Proportion of the mean genetic value of the twenty selected phenotypes that was due to the individual genetic value relative to the family genetic value for each of the three sites with $h^2=.2$ and 30 trees per family.

Design	Dierks			Cloudy			Hempwallace		
	Rep	Nay	Plot	Rep	Nay	Plot	Rep	Nay	Plot
10-tree row plots	0.61	0.62	0.60	0.40	0.41	0.39	0.43	0.43	0.41
5-tree row plots	0.58	0.59	0.55	0.39	0.41	0.38	0.43	0.42	0.38
noncontiguous plots	0.54	0.55	0.52	0.38	0.39	0.35	0.42	0.42	0.35
2-tree plots	0.55	0.56	0.46	0.39	0.39	0.30	0.43	0.42	0.33
single-tree plots	0.55	0.55	-	0.38	0.39	-	0.43	0.42	-

Table 14. Proportion of the mean genetic value of the twenty selected phenotypes that was due to the individual genetic value relative to the family genetic value for the Cloudy plantation with $h^2=.2$ and $h^2=.3$; 50 trees per family and 30 families.

Design	$h^2=.2$			$h^2=.3$		
	Rep	Nay	Plot	Rep	Nay	Plot
25-tree block plots	0.59	0.59	0.58	0.60	0.60	0.60
10-tree row plots	0.49	0.50	0.46	0.52	0.53	0.48
5-tree row plots	0.50	0.49	0.45	0.53	0.52	0.47
noncontiguous plots	0.49	0.49	0.43	0.53	0.53	0.44
2-tree plots	0.51	0.50	0.40	0.54	0.52	0.44
single-tree plots	0.51	0.50	-	0.54	0.53	-

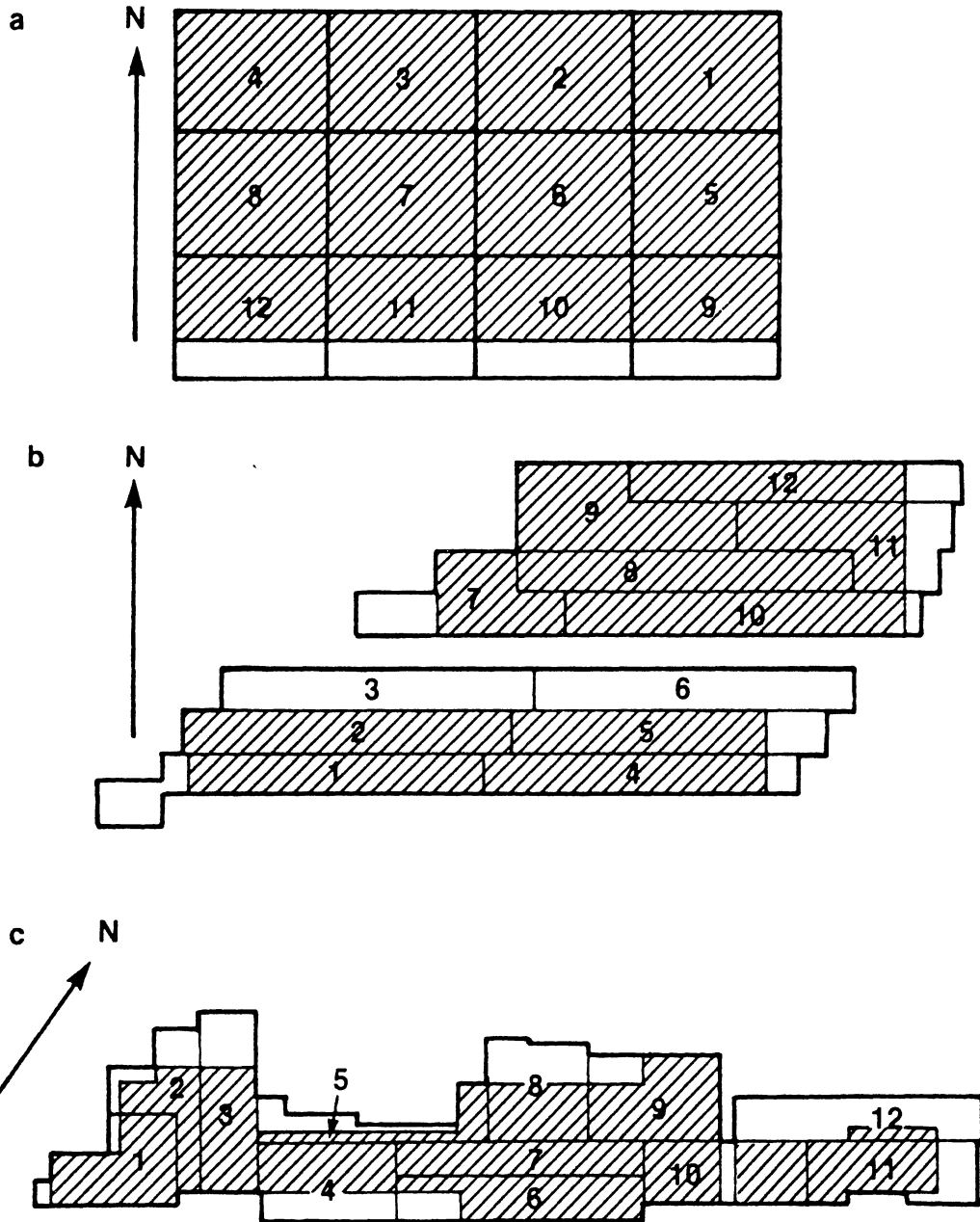Fig. 1. Plantations used in the simulation model; AOCG7505, located near Dierks, Arkansas (a), AOWF8088, located near Cloudy, Oklahoma (b), and AOWF8085, located near Hempwallace, Arkansas (c).

Fig. 2. Mean percentage of maximum genetic values plotted over number of trees per plot for each of three selection systems. Values were averaged over the three sites with a heritability of .20 and 50 or 60 trees per family.
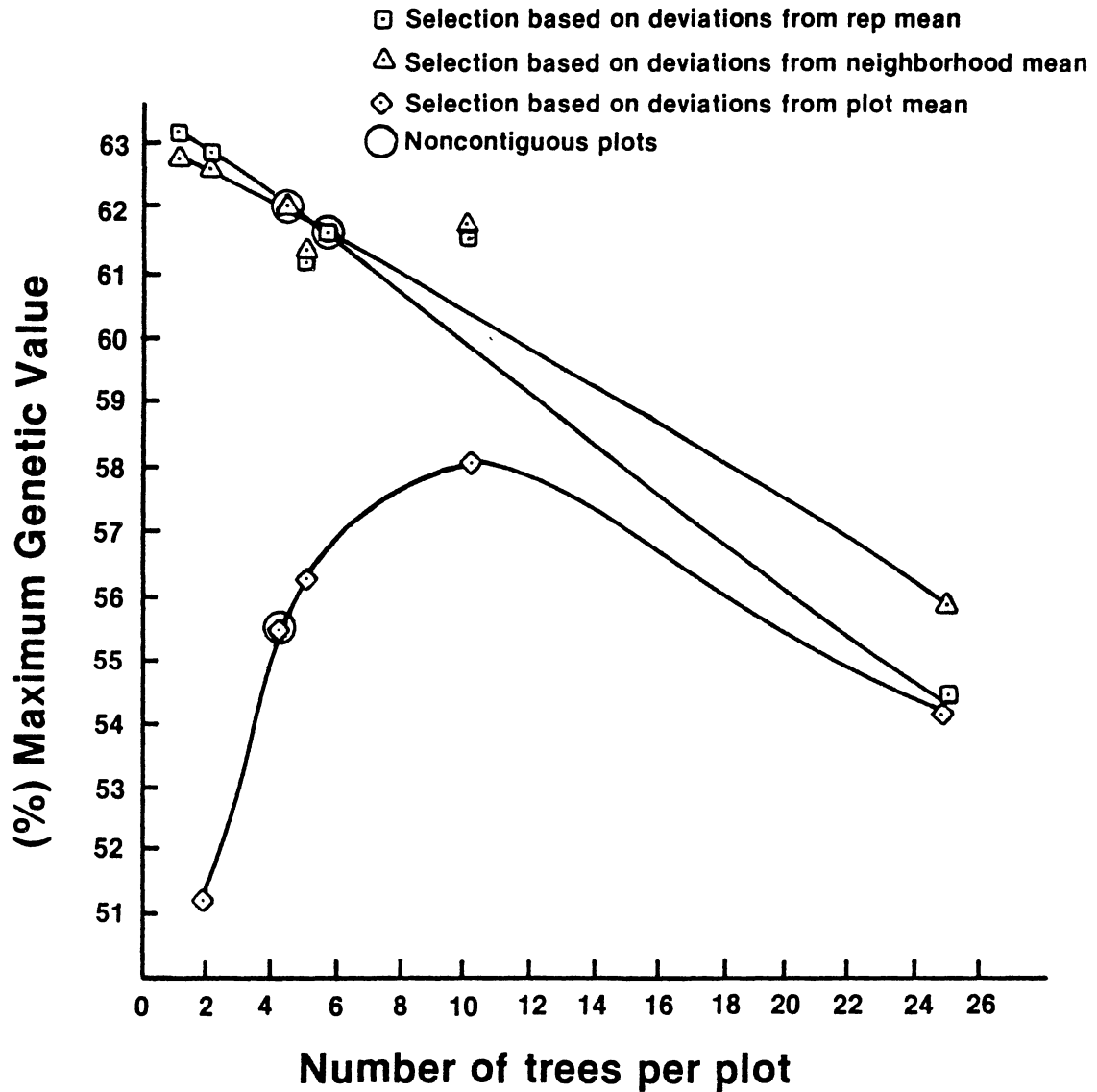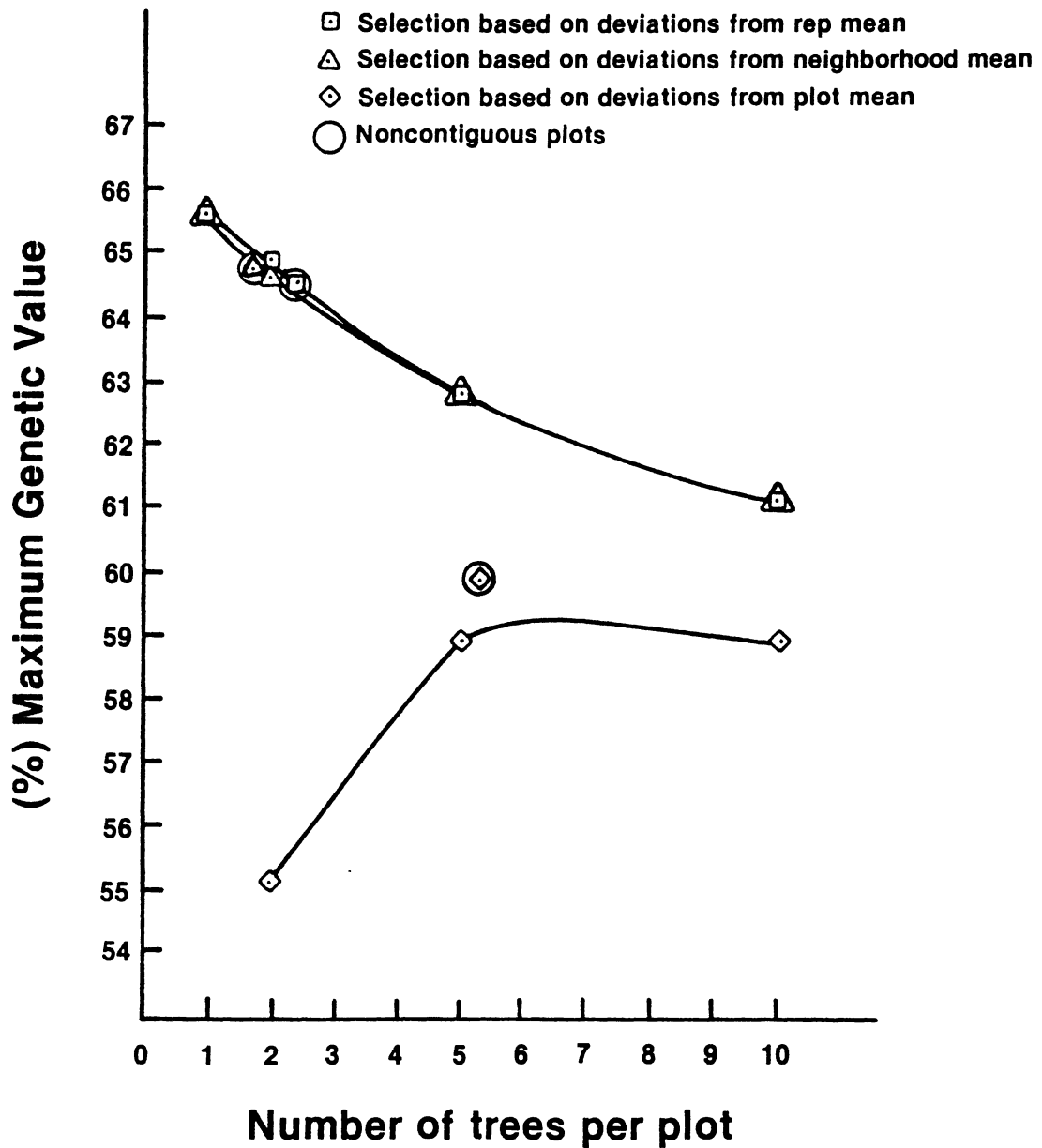
Fig. 3. Mean percentage of maximum genetic values plotted over number of trees per plot for each of three selection systems. Values were averaged over the three sites with a heritability of .20 and 30 trees per family.

PART III

NEIGHBORHOOD COMPARISON AS A BASIS FOR

SELECTION IN LOBLOLLY PINE

(PINUS TAEDA L.)

SUMMARY

Optimizing within-family selection in a progeny test may considerably improve genetic gains. A comparison of potential select trees with unrelated neighbors which occupy an area within the range of significant positive environmental correlations is suggested as an optimal selection method for individual trees within families. A selection index is described which includes neighborhood data in addition to family and individual tree information. Construction of semi-variograms for three loblolly pine (Pinus taeda L.) progeny test sites in southeastern Oklahoma and southwestern Arkansas indicates that the size and shape of the optimum neighborhood area for selection is site specific, depending upon spatial dependency relationships within the site.

Key words: Neighborhood selection, semi-variogram, selection index

INTRODUCTION

Selection of parents for testing and breeding of forest
trees increasingly is being carried out in progeny tests.
Trees in such tests have at least one known parent, are
even-aged, and have already benefited from some degree of
genetic improvement. Thus, selection of superior families
and of superior individuals within families is an important
aspect of progeny testing. The primary purpose of most
progeny tests, to evaluate seed orchard trees as potential
parents of commercial seed, necessitates locating the tests
on sites representative of the range of intended commercial
site types. Commercial loblolly pine (_Pinus_ _taeda_ L.)
sites in southeastern Oklahoma and southwestern Arkansas
often contain a high degree of variability due to factors
such as changes in slope and aspect, soil texture, soil
moisture, and fertility. Consequently, progeny tests are
often located on highly heterogeneous sites.

Some combination of family and within-family
information is almost always used in selection of forest
trees from progeny tests. The progeny test design
determines the efficiency of estimation of family means
(Barnes and Schweppenhauser 1979). Even on a highly
variable site, families can be accurately ranked provided

an efficient test design is used (Loo-Dinkins, Part I).

Heritabilities of growth traits in forest trees are

generally low to moderate (Zobel and Talbert 1984); thus,

given the same selection intensity, selection at the family

level is typically more effective than within-family

selection (Falconer 1981). Because of large family size,

within-family selection intensity is often considerably

higher than the intensity of among family selection. This

would allow within-family selections to contribute

significantly to genetic gain if an accurate selection

method were employed. Thus the selection system that will

most accurately identify the best individuals within

families is of great importance in a breeding program.

The optimum within-family selection method may depend

upon the size and shape of the area in which comparisons

are made relative to the type of variation present in the

site. Commonly, trees are compared with family plot or

replicate means. Particularly on a site with a gradient,

an effective family plot should sample the range of

microsites within a replicate, but the environmental

variability within a comparison area for individual tree

selection should be minimized. Because of this conflict

and the usual small size of family plots, within-family

selection based on deviations from plot means may be

inaccurate, particularly in variable environments (Loo-

Dinkins, Part II). Progeny tests usually include large

numbers of families (often 50 to 75). The most common test

designs consist of replicates of 4- to 6-tree family row plots; thus, replicate size may be large when compared to the area in which positions may be environmentally correlated. Neighborhoods, based on the size of areas where all trees are significantly, positively correlated with each other due to environmental conditions, may constitute the optimal comparison area size.

Gardner (1961) developed a selection method for corn (Zea mays L.) which reduced the environmental variance within a selection area by superimposing a grid on a test field and selecting a set percentage of plants from within each stratum. This allowed greater separation of genetic and environmental effects, thus more effective selection. In loblolly pine progeny tests, the number of selections within a neighborhood would not be expected to be constant. Typically, some percentage of the best families and only one or two individuals within those families are selected to minimize inbreeding. Selection using deviations from a neighborhood mean or making use of environmental correlations in an index may result in gains greater than those attained by other commonly employed selection methods due to better estimates of individual genetic values.

Optimal neighborhoods for selection would be expected to be site specific. The nature of the site variability may be characterized using regionalized variable theory (Burgess and Webster 1980, Vieira et al. 1981). Burgess and Webster (1980) described a function called a semi-variogram

which provides a measure of the average similarity between points separated by a given distance. Semi-variograms are used to describe the spatial dependence relationships within a sampled region. A measure of the average similarity between environmental values associated with each position in a plantation may be plotted over the distance between sets of positions to determine the average size and shape of a region with spatial dependency among environmental values. Application of this theory to the nature of site variability within specific progeny tests could be very useful in determining the optimal size and shape of neighborhoods for selection.

The objective of this study was to characterize the environmental tree to tree variation within three different site types in southeastern Oklahoma and southwestern Arkansas and to develop a selection index which includes environmental correlations as a source of information.

MATERIALS AND METHODS

Three Weyerhaeuser Company loblolly pine progeny tests
provided an environmental data base for site
characterization. The Dierks plantation, AOGC7505, is a
half-sib progeny test located near Dierks, Arkansas. The
site was considered the most uniform of the three with no
apparent gradient or patchiness. The test is rectangular
with 12 replicates, each consisting of a 15 X 18 tree block
with 54 families. Eight of the replicates consist of 5-
tree row plots. The remaining four replicates contain 5-
tree noncontiguous plots. The Cloudy plantation, AOWF8088,
is a full-sib progeny test located near Cloudy, in
southeastern Oklahoma. This plantation was considered to
be the most variable of the three sites. The site
straddles a ridge oriented in an east-west direction and
six replicates of 34 5-tree row family plots slope downward
on each side of the ridge. The Hempwallace plantation,
AOWF8085, is a full-sib progeny test located near
Hempwallace in southwestern Arkansas. The plantation is on
a wet site, causing a pattern of patchy variability. The
test consists of 40 families arranged in 12 replicates.
Eight of the replicates have a 5-tree row plot arrangement,
and the remaining four consist of 5-tree noncontiguous

plots.

Residual values for height were obtained for each position in each plantation by subtracting the plantation mean and the family effect from the height value of each tree. Eight-year data from the Dierks plantation and five-year data from the other two plantations were used. A component of the genetic effect remained in the residual values. However, the environmental effect was assumed to be sufficiently larger than the remaining genetic effect to allow the residual values to primarily reflect the microsite values, as suggested by Franklin (1971). In the case of full-sib families, subtraction of the family effect removes one-half of the additive genetic effect. For example, if individual tree heritability is equal to 0.2, approximately 90% of the variance of the residual values would be due to nonadditive genetic and environmental causes. The residual values became the environmental map for each plantation, and each position was given a row and column number.

Each position, excluding the outer two rows and columns in each plantation, was considered to be the center of a 25-tree square neighborhood. Positions were numbered consecutively from the upper left-hand corner of the neighborhood (Fig. 1) so that position 13 was always the center of the moving neighborhood. The average correlation between each pair of positions within the 25-tree moving neighborhood was estimated for each plantation. This

provided estimates of the environmental correlations between each pair of microsites within a 5-tree proximity in any direction.

A selection index was derived that made use of environmental correlations among neighborhood trees as well as individual tree and family mean data. All trees within a neighborhood were assumed to be unrelated; thus, the index would be appropriate only for single-tree or noncontiguous plot designs. The selection index including neighborhood information was compared with a selection index constructed with individual and family information alone, using a simulation model. Neighborhood size was nine trees. Values were generated to correspond to a heritability of 0.20 to 0.40 and a phenotypic, environmental correlation among neighbors of 0.15 to 0.35. Both indices were applied to the populations with a selection intensity of 0.05.

To further characterize the site variability in each plantation, a semi-variogram was constructed to examine the relationship between spatial dependency and distance between trees. The semi-variance function is:

$$\Gamma(h) = (1/2 \ m) \ E \ \{[z(x_i)-z(x_i-h)]^2\}$$

where $\Gamma(h)$ is the semi-variance between pairs of positions, m is the number of pairs of positions included in the summation, $z(x_i)$ is the height residual at position i, and $z(x_i-h)$ is the height residual value at a distance h from position i. The semi-variance reflects the degree to which

two points at a given distance apart are correlated. The
relationship between a semi-variance and the correlation
between two points may be expressed as:

$$\Gamma(h) = \sigma_p^2 (1-r_{ij})$$

where σ_p^2 is the phenotypic variance and r_{ij} is the
correlation between two points (i and j). A $\Gamma(h)$ value was
estimated for pairs of trees at distances of one to nine
positions apart within rows and columns for the Hempwallace
and Cloudy plantations and at distances of one to 14
positions apart for the Dierks plantation. $\Gamma(h)$ values were
plotted over distance both in the direction of rows and in
the direction of columns to construct a semi-variogram.

RESULTS AND DISCUSSION

The variance of residual values for height included a portion of the genetic variance; one-half of the additive genetic variance in the case of the full-sib plantations, and three-quarters of the additive genetic variance in the half-sib plantation. The family plots fell within columns. Thus, comparisons of correlation estimates across rows vs. across columns (Table 1) provide an indication of whether error in estimating family effects influenced the correlations among positions. Correlations within family plots (position 1 with positions 6, 11, 16, and 21) are slightly higher than correlations across family plots (position 1 with positions 2, 3, 4, and 5) on the Dierks plantation. The opposite trend is apparent in each of the other plantations. Thus, the assumption that residual values primarily reflect microsite quality appears valid.

The Dierks plantation was considered to be the most uniform with existing variability appearing random. Correlations between positions were lower in general (r = 0.11 to 0.35) at this site than at either of the other sites (r = 0.22 to 0.45 at Hempwallace and r = 0.19 to 0.63 at Cloudy). The Hempwallace plantation appears to have patchy variability, and the size of the patches is probably

70

larger than the 25-tree neighborhoods. The correlation
estimates do not differ greatly between rows and columns,
and remain relatively high at a distance of four positions
apart. The gradient effect is apparent in the correlation
estimates for the Cloudy site. The correlations are high,
r = 0.54 to 0.63, between positions 1 to 5 within rows;
perpendicular to the gradient. The magnitude of
correlations declines rapidly with distance parallel to the
gradient, however, r = 0.55 to 0.20, for positions 1 with 2
and 1 with 5, respectively.

Given the magnitude of the correlation coefficients
between trees, particularly on the more variable sites,
environmental covariances should be useful in the selection
process. Increased gains should result from comparison of
potential select trees in a progeny test with unrelated
neighbors located in positions expected to be significantly
positively correlated with the position occupied by the
potential selected tree. Selection based on deviations
from a neighborhood mean was included in a simulation study
to compare various progeny test designs in combination with
within family selection methods (Loo-Dinkins, Part II) She
found that neighborhood selection was always better than
selection based on deviations from a plot mean and was as
good as selection based on deviations from replicate
means. With large replicate size, selection on the basis
of deviations from a neighborhood mean generally performed
better than selection based on deviations from replicate

means.

Selection Index

One approach to using neighborhood information is by inclusion of neighborhood covariances as an additional source of information in a selection index. An index was constructed using individual tree value, family mean, and neighborhood mean. The form of the index is:

$$I = P_1 b_1 + P_2 b_2 + P_3 b_3,$$

where I is the index value, P_1 is the individual tree value, P_2 is the family mean, P_3 is the neighborhood mean, and b_1, b_2, and b_3 are the weighting coefficients. The coefficients are found by solving the set of equations:

$$b_1 P_{11} + b_2 P_{12} + b_3 P_{13} = A_{11}$$

$$b_1 P_{21} + b_2 P_{22} + b_3 P_{23} = A_{12}$$

$$b_1 P_{31} + b_2 P_{32} + b_3 P_{33} = A_{13}$$

where P_{11} is the phenotypic variance, P_{12} is the phenotypic covariance between individual and family mean, P_{13} is the phenotypic correlation between the individual and the neighborhood mean, $P_{21} = P_{12}$, P_{22} is the variance of a family mean, P_{23} is the covariance between the family and neighborhood means, $P_{31} = P_{13}$, $P_{32} = P_{23}$, P_{33} is the variance of a neighborhood mean, A_{11} is the covariance between the individual's phenotype and its breeding value, A_{12} is the covariance between the individual's breeding value and the family mean, and A_{13} is the covariance between the individual breeding value and the neighborhood mean.

The variance of a family mean can be expressed as:

$$\{[1+t_f(N-1)]/N\}\sigma_P^2.$$

where t_f is the phenotypic correlation among family members, and N is the number of individuals within a family. The additive genetic component of the variance of a family mean is:

$$\{[1+r(N-1)]/N\}\sigma_A^2$$

where r is the correlation among individuals within a family, and σ_A^2 is the additive genetic variance (Falconer 1981). The variance of a neighborhood mean is:

$$\{[1+t_n(n-1)]/n\}\sigma_P^2$$

where n is the number of trees within a neighborhood, and t_n is the phenotypic, environmental correlation among neighbors. Thus:

$$P_{12} = \{[1+t_f(N-1)]/N\}\sigma_P^2$$

$$P_{13} = \{[1+t_n(n-1)]/n\}\sigma_P^2$$

$$P_{22} = P_{12}$$

$$P_{23} = \{[1+t_n(n-1)+t_f(N-1)]/Nn\}\sigma_P^2$$

$$P_{33} = P_{13}$$

$$A_{11} = \sigma_A^2$$

$$A_{12} = \{[1+r(N-1)]/N\}\sigma_A^2$$

$$A_{13} = (1/n)\sigma_A^2$$

Letting: $[1+t_f(N-1)]/N = F,$

$[1+t_n(n-1)]/n = H,$

$[1+t_n(n-1) + t_f(N-1)]/Nn = C,$ and

$[1+r(N-1)]/N = R,$

$$b_1 = h^2([FH-C^2-RH(F-C)+(1/N)F(C-H)] / [FH(1+2C-F-H)-C^2])$$

$$b_2 = h^2([-H(F-C)+RH(1-H)+(1/N)(FH-C)] / [FH(1+2C-F-H)-C^2])$$

$$b_3 = h^2([F(C-H)+R(FH-C)+(1/N)F(1-F)] / [FH(1+2C-F-H)-C^2])$$

The coefficients were programmed with various
combinations of t_n, t_f, N, n, σ_P^2, and σ_A^2 to examine the
relative magnitudes of the coefficients under different
conditions. When trees within neighborhoods are positively
correlated, b_3 will always be negative; and if t_n is
approximately equal to the heritability (h^2), the magnitude
of b_1 is approximately the same as $-b_3$. In comparisons
between genotypes selected using an index with and without
the neighborhood information, on simulated data with
varying h^2 and t_n values, the selection index using
neighborhood information was consistently better by
approximately 4% when t_n was approximately equal to 0.30
(Table 2). If the average breeding value for height of
selected trees using an index which includes only family
and individual tree information, is 2.5 feet, representing
one-sixth of the total height (15 feet), an increase of 4%
due to including neighborhood information, would increase
the average genetic value to 2.6. This represents an
increase in total height of 1.1%. The heritability was
initially varied from 0.20 to 0.40 but it affected both
indices in the same way for a given t_n value.

The most serious drawback to adding new parameters to

selection would include positions within a region showing spatial dependency. Thus the semi-variance between any two positions in a neighborhood should be less than the total environmental variance in the plantation. If the semi-variogram plot has a similar shape and semi-variances are similar in magnitude in both directions, circular neighborhoods with a maximum diameter equal to the range of the semi-variogram would be optimal. If square neighborhoods are used, the maximum distance between diagonal corners should equal the range of the semi-variogram. The Hempwallace semi-variogram plot shows similar curves for each direction although rows appear to have slightly smaller semi-variances among positions than columns (Fig. 2). A circular or square neighborhood with a maximum diameter equal to a distance of about eight or nine positions should be optimal.

If semi-variances are consistently smaller in one direction, neighborhoods should be oblong or rectangular with the long axis in the direction with smallest semi-variances. The maximum length of the short axis should not exceed the range of the semi-variogram with largest semi-variances. Optimum neighborhoods in the Dierks plantation would be oriented with the long axis within rows, not exceeding a distance of about 11 positions, and the short axis, in the direction of columns, would not exceed a distance of approximately six positions (Fig. 1).

Optimum neighborhoods indicated by the semi-variogram

plot for the Cloudy site would include a maximum distance of three positions within columns (Fig. 3). The semi-variogram for the within rows direction is almost flat and the semi-variance for an average distance of nine positions apart within rows is less than the average semi-variance between any two positions within columns. Thus a long, narrow neighborhood should be optimal.

CONCLUSIONS

A measure of environmental correlations or spatial
dependency within a neighborhood should allow more accurate
within family selection on a variable site. The inclusion
of neighborhood information in a selection index can
increase efficiency relative to an index including only
family mean and individual values by approximately 4% when
the phenotypic correlation within neighborhoods is
approximately 0.30. Among the three study sites, the
overall mean correlation within a 25-tree square
neighborhood was approximately 0.30. Thus, on the average,
with square neighborhoods, the use of neighborhood
information in an index would increase the genetic gain
slightly. The Cloudy site had considerably higher
correlations than the other two sites. Thus, the
neighborhood approach may be more valuable in a highly
variable test site.

The size and shape of optimal neighborhoods for
selection is site specific depending upon the spatial
dependence relationships within the site. A semi-variogram
is a useful tool in characterizing the nature of site
variability in a plantation. It could be used easily to
determine the size and shape of an area with sufficiently

high tree-to-tree environmental correlations to merit a

neighborhood selection approach.

REFERENCES

Barnes RD, Schweppenhauser MA (1979) Genetic control
of 1,5-year-old traits in _Pinus patula_ Schiede et
Deppe and a comparison of progeny test methods. Silvae
Genet 28:156-167

Burgess TM, Webster R (1980) Optimal interpolation
and isarithmic mapping of soil properties. I. The semi-
variogram and punctual kriging. J Soil Sci 31:315-
331

Falconer DS (1981) Introduction to quantitative genetics.
2nd ed Longman Inc, New York

Franklin EC (1971) Statistical validity of single-tree
plot in forest genetics research. Silvae Genet
20:73-75

Gardner CO (1961) An evaluation of effects of mass
selection and seed irradiation with thermal neutrons on
yield of corn. Crop Sci 1:241-245

Loo-Dinkins JA Comparisons among combinations of progeny
    test designs and selection systems for maximum genetic
    gain in loblolly pine (*Pinus taeda* L.) Part I. Unpubl.
    Ph.D. dissertation. OSU, Stillwater, OK

Loo-Dinkins JA Comparisons among combinations of progeny
    test designs and selection systems for maximum genetic
    gain in loblolly pine (*Pinus taeda* L.) Part II. Unpubl.
    Ph.D. dissertation. OSU, Stillwater, OK

Vieira SR, Nielsen DR, Biggar JW (1981) Spatial
    variability of field-measured infiltration rate. Soil
    Sci Soc Amer J 45:1040-1048

Table 1. Average Pearson correlation estimates between pairs of residual values for tree height in moving 25-tree square neighborhoods for each of three plantations.

aPositions correlated	plantation		
	Dierks	Hempwallace	Cloudy
1-2	0.25	0.45	0.63
1-3	0.21	0.40	0.57
1-4	0.20	0.36	0.55
1-5	0.18	0.35	0.54
1-6	0.35	0.38	0.55
1-7	0.22	0.32	0.51
1-8	0.22	0.33	0.47
1-9	0.16	0.35	0.48
1-10	0.15	0.34	0.50
1-11	0.27	0.34	0.37
1-12	0.19	0.36	0.42
1-13	0.21	0.33	0.40
1-14	0.17	0.34	0.37
1-15	0.18	0.30	0.39
1-16	0.25	0.30	0.22
1-17	0.17	0.29	0.22
1-18	0.15	0.27	0.27
1-19	0.13	0.27	0.23
1-20	0.13	0.30	0.23
1-21	0.23	0.22	0.20
1-22	0.18	0.26	0.18
1-23	0.12	0.30	0.19
1-24	0.11	0.29	0.21
1-25	0.11	0.25	0.19

aAll correlations are statistically significant at P=0.01.

Table 2. Ratio of simulated genetic values selected with a neighborhood index to values selected using only family and individual information. Five levels of neighborhood phenotypic correlations were tested.

Ratio	Neighborhood intraclass correlation				
	0.15	0.20	0.25	0.30	0.35
Index1/Index2	1.01	1.01	1.02	1.04	1.04

a Results are the mean of eighty iterations.

```
                    ←—columns—→
            1      6     11    16    21

        ↑   2      7     12    17    22
        |
  rows   3      8     13    18    23
        |
        ↓   4      9     14    19    24

            5     10     15    20    25
```

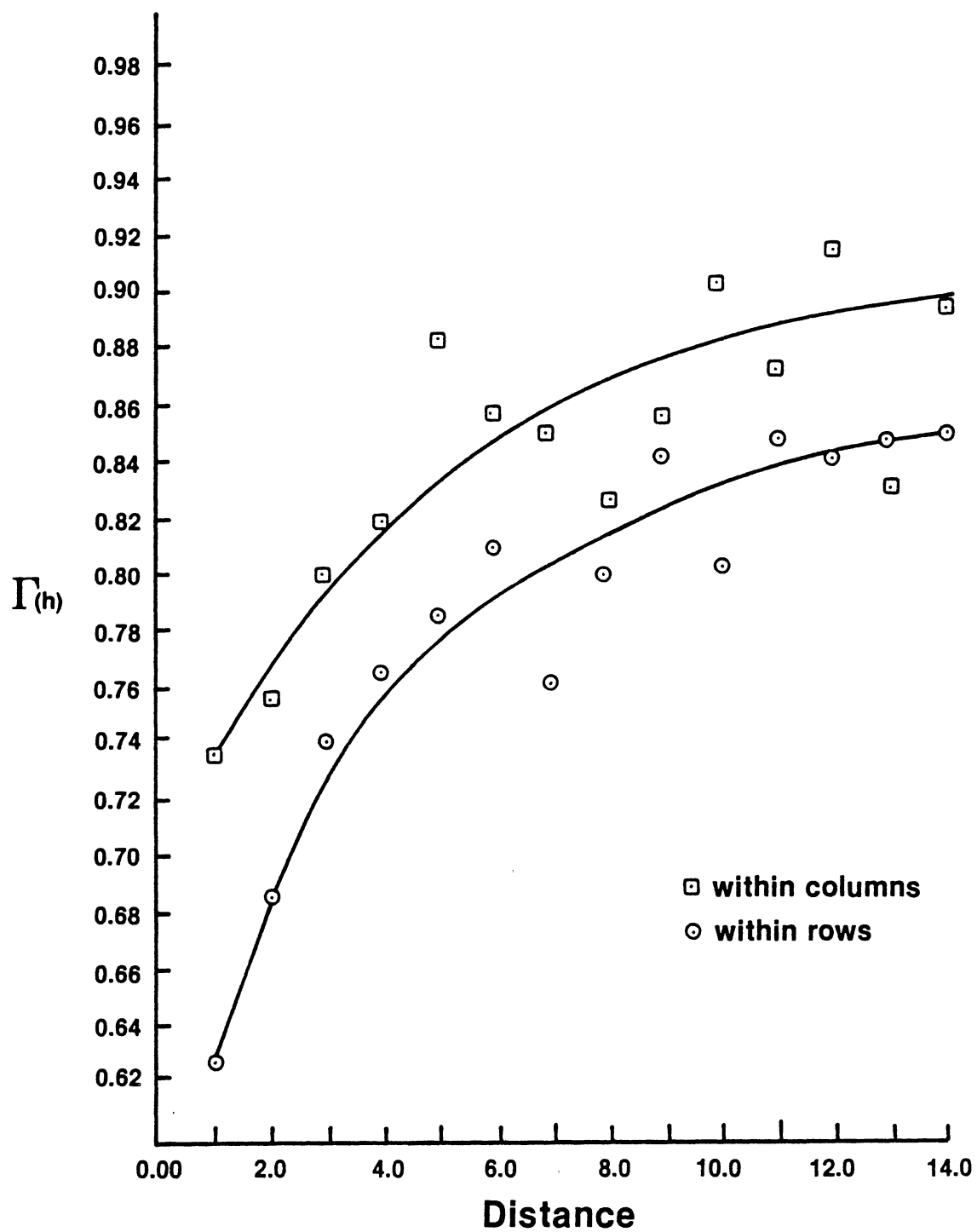Fig. 1. The numbering of positions within the moving neighborhood.

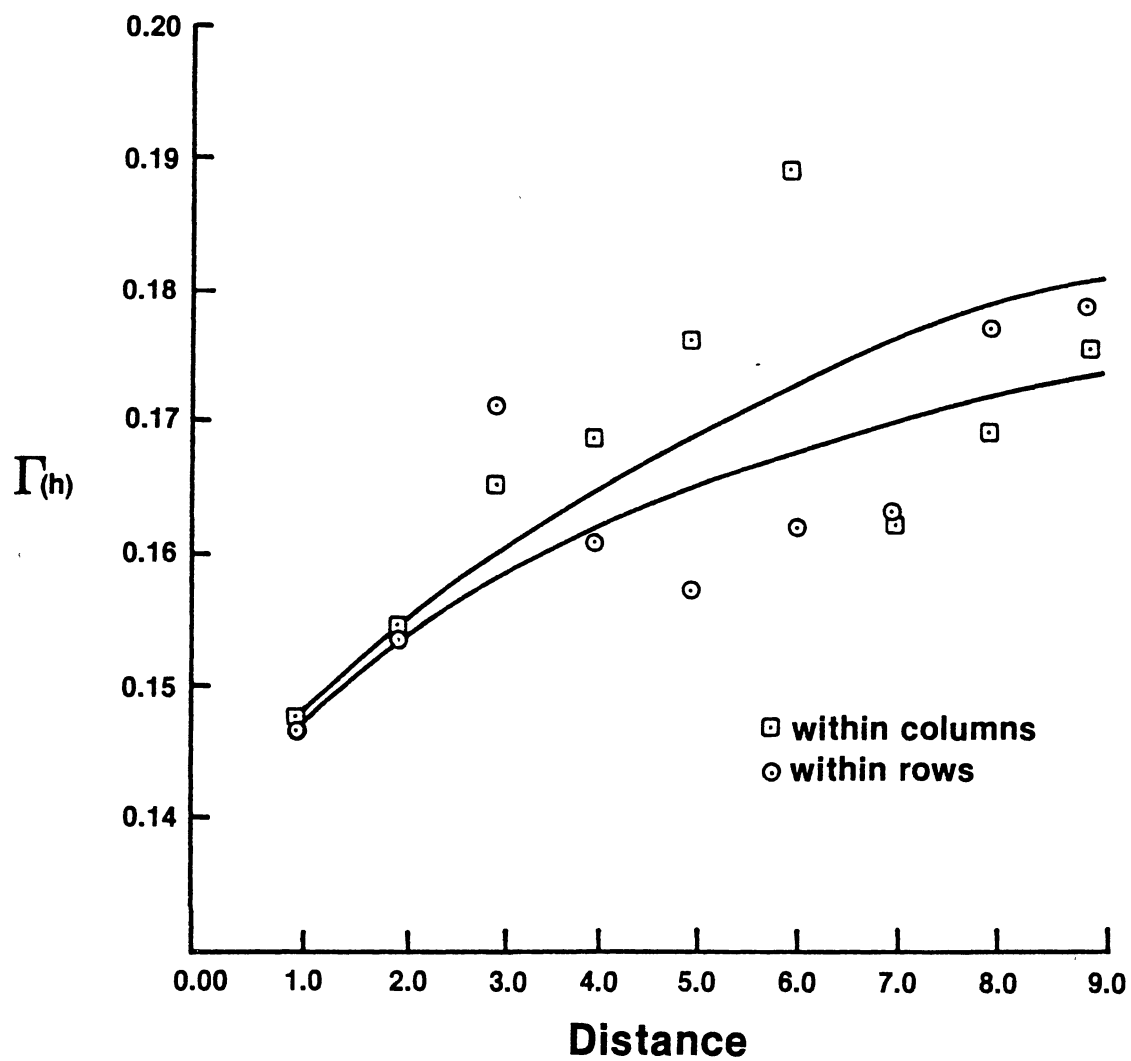Fig. 2. Semi-variogram for the Dierks plantation in two directions.

Fig. 3. Semi-variogram for the Hempwallace plantation in two directions.
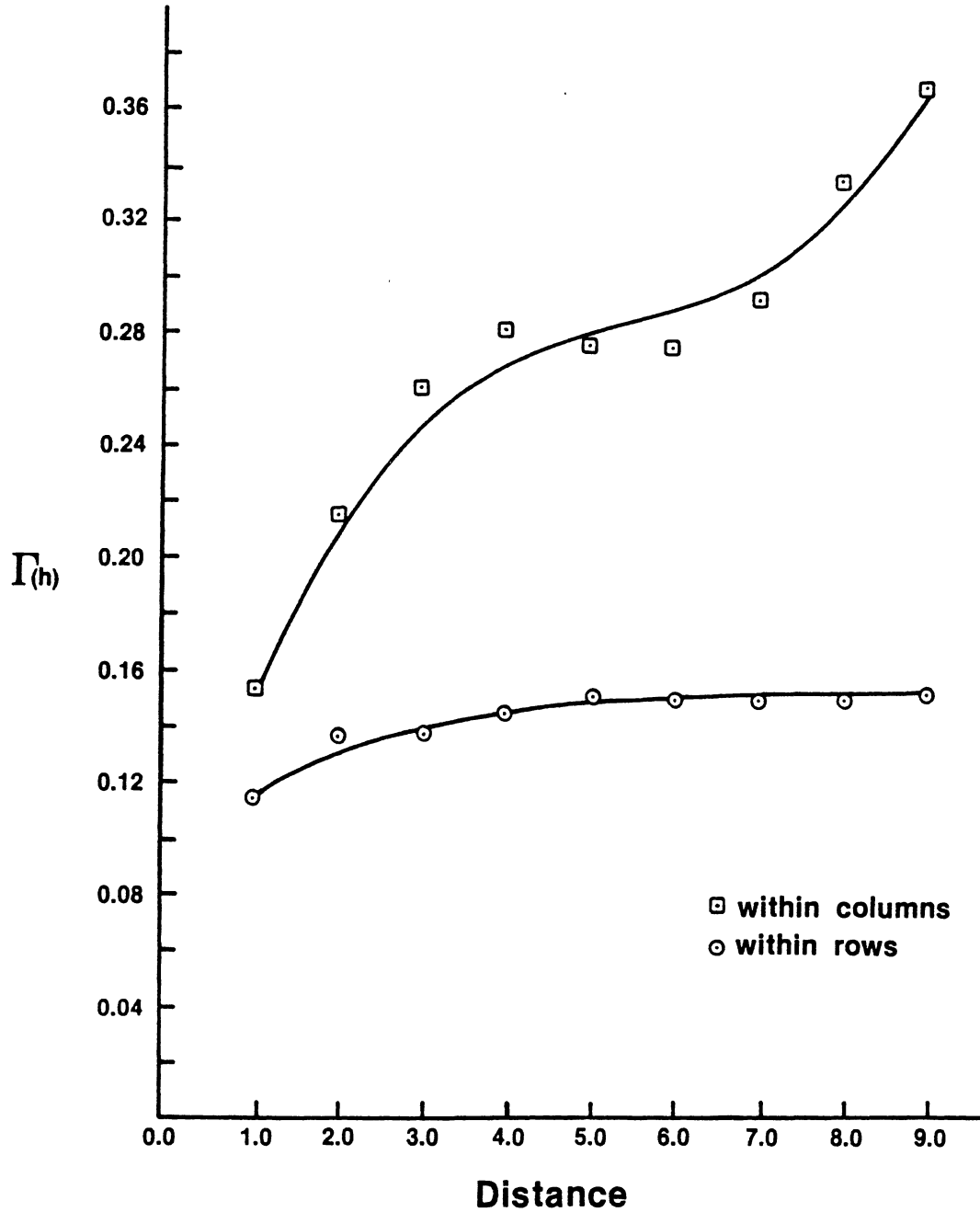
Fig. 4. Semi-variogram for the Cloudy plantation in two directions.

APPENDIX A

COMPUTER MODEL USED IN DESIGN AND

SELECTION SYSTEM SIMULATIONS

(SAS)

```
1      OPTIONS DQUOTE;
2      %MACRO GENET(N, NFAM1, NFAM2, NTREE1, NTREE2, TPOS);
3      %DO I=1 %TO &N;
4      %* N=THE NUMBER OF ITERATIONS;
5      DATA GEN2;
6      *CREATING THE GENETIC POPULATION FOR THE FIRST DESIGN;
7      DO FAM=1 TO &NFAM2;
8      GF=RANNOR(0)*.2537;
9      DO TREE=1 TO &NTREE2;
10     GT=RANNOR(0)*.4394;
11     GEN=GF+GT;
12     OUTPUT;
13     END;
14     END;
15     RUN;
16     PROC SORT DATA=GEN2 OUT=STANDARD;
17     BY DESCENDING GF DESCENDING GT;
18     RUN;
19     DATA ONE;
20     DO F=1 TO 10;
21     DO T=1 TO &NTREE2;
22     OUTPUT; END;
23     END;
24     RUN;
25     DATA STANDARD; MERGE ONE STANDARD;
26     RUN;
27     DATA STANDARD; SET STANDARD;
28       IF T< 3 AND .<F<11;
29     RUN;
30     PROC MEANS NOPRINT DATA=STANDARD;
31     OUTPUT OUT=STANDARD MEAN=BESTGEN;
32     VAR GEN GF GT;
33     RUN;
34     PROC PRINT;
35     TITLE "BEST GENOTYPES FOR ITERATION &I";
36     RUN;
37     %MACRO DESIGN(NDES, NFAM1, NFAM2, TPOS);
38     %DO DES=3 %TO &NDES;
39     %*NDES=THE NUMBER OF DESIGNS TO BE TESTED;
40     DATA FRAME;
41       %IF &DES>1 %THEN %GOTO ST2;
42     %ELSE %IF &DES=1 %THEN %DO;
43     %LET NFAM=&NFAM1;
44     %LET NTREE=&NTREE1;
45     %LET NSEL=%EVAL(&NTREE.*10);
46     %END;
47     %GOTO START;
48       %ST2: %IF &DES>1 %THEN %DO;
49     %LET NFAM=&NFAM2;
```

```
50          %LET NTREE=&NTREE2;
51         .%LET NSEL=%EVAL(&NTREE.*10);
52          %END;
53          %START:
54          %IF &DES>1 %THEN %GOTO LO;
55          %ELSE %IF &DES=1 %THEN %DO;
56          %*DESIGN 1=6*10 BLOCK PLOTS;
57           P=60;
58           %LET Q=60;
59           C=40;
60           S1=10;
61           S2=6;
62          %END;
63          %GOTO L7;
64          %LO: %IF &DES>2 %THEN %GOTO L1;
65          %IF &DES=2 %THEN %DO;
66          %*DESIGN 2 = 5*6 BLOCK PLOTS;
67           P=30;
68           %LET Q=30;
69           C=40;
70           S1=5;
71           S2=6;
72          %END;
73          %GOTO L7;
74          %L1: %IF &DES>3 %THEN %GOTO L3;
75          %IF &DES=3 %THEN %DO;
76          %*DESIGN 3 = 5-TREE NONCONTIGUOUS PLOTS;
77           P=5;
78           %LET Q=5;
79           C=40;
80           S1=20;
81           S2=24;
82          %END;
83          %GOTO L7;
84          %L3: %IF &DES>4 %THEN %GOTO L4;
85          %IF &DES=4 %THEN %DO;
86          %*DESIGN 4 = 5 TREE ROW PLOTS;
87           P=5;
88           %LET Q=5;
89           C=40;
90           S1=20;
91           S2=24;
92          %END;
93          %GOTO L7;
94          %L4: %IF &DES>5 %THEN %GOTO L5;
95          %IF &DES=5 %THEN %DO;
96          %*DESIGN 5 = 10 TREE ROW PLOTS;
97           P=10;
98           %LET Q=10;
99           C=40;
100          S1=40;
101          S2=24;
102         %END;
103         %GOTO L7;
104         %L5: %IF &DES>6 %THEN %GOTO L6;
105         %IF &DES=6 %THEN %DO;
106         %* DESIGN 6 = 2 TREE PLOTS;
107          P=2;
108          %LET Q=2;
109         C=40;
110         S1=8;
111         S2=24;
112         %END;
113         %GOTO L7;
114         %*DESIGN 7 = SINGLE TREE PLOTS;
115         %L6:   P=1;
116          %LET Q=1;
117         C=40;
118         S1=8;
119         S2=12;
120         %L7:
121         %LET NPOS=%EVAL(&Q.*&NFAM.);
122         %LET NREP=%EVAL(&TPOS./&NPOS.);
```

```
123        QN=&NPOS;
124       NR=&NREP;
125       PN=P*&NFAM;
126        B=1;
127        J=1;
128        K=0;
129        M1=1;
130       POS=0;
131       REP=0;
132       %IF  &DES=2 %THEN %DO;
133        M2=S1;
134       COS=C/S1;
135       CN=&NPOS./COS;
136       DO REP=1 TO &NREP;
137       DO V1=B TO CN BY P;
138       DO BL=1 TO COS;
139       U=RANUNI(0)*100;
140       DO V2=J TO S2;
141       DO V3=M1 TO M2;
142       POS=V3;
143       OUTPUT;
144       END;
145        M1=M1+C;
146       M2=M2+C;
147       END;
148       M1=M1+S1-(C*S2);
149       M2=M1-1+S1;
150       END;
151       M1=M1+C*(S2-1);
152       M2=M1+S1-1;
153       END;
154       END;
155       %END;
156       %ELSE %IF &DES=3   %THEN %DO;
157       K=P;
158       M2=S1;
159       COS=C/S1;
160       RR=NR/COS;
161       DO R=1 TO RR;
162       DO V=1 TO COS;
163       REP=REP+1;
164       DO V1=1 TO S2;
165       DO V2=M1 TO M2;
166       U=RANUNI(0)*100;
167       POS=V2;
168       OUTPUT;END;
169       M1=M1+C;
170       M2=M1+S1-1;
171       END;
172       M1=M1+S1-(C*S2);
173       M2=M1+S1-1;
174       END;
175       M1=M1+C*(S2-1);
176       M2=M1+S1-1;
177       END;
178       %END;
179       %ELSE %IF &DES>3 %THEN %DO;
180       K=P;
181       M2=S1;
182       COS=C/S1;
183       RR=NR/COS;
184       DO R=1 TO RR;
185       DO V=1 TO COS;  .
186       REP=REP+1;
187       DO V1=1 TO S2;
188       DO V2=M1 TO M2 BY P;
189       U=RANUNI(0);
190       DO V3=J TO K;
191       POS=V3;
192       OUTPUT;  END;
193       J=J+P;
194       K=K+P;
195       END;
```

```
196        M1=M1+C;
197        M2=M1+S1-1;
198        J=M1;
199        K=J+P-1;
200        END;
201        M1=M1+S1-(C*S2);
202        M2=M1+S1-1;
203        J=M1;
204        K=J+P-1;
205        END;
206        M1=M1+C*(S2-1);
207        M2=M1+S1-1;
208        J=M1;
209        K=J+P-1;
210        END;
211        %END;
212        KEEP POS REP U;
213        RUN;
214        PROC SORT OUT=FRAME;
215        BY REP U;
216        RUN;
217        DATA IDEN;
218         NF=&NFAM2;
219        DO X=1 TO NF;
220         A=1;
221        DO Y=1 TO &Q;
222        ID=0;
223        DO S=A TO &NREP;
224        ID=ID+1;
225        OUTPUT;
226        END;
227        END;
228        END;
229        KEEP ID;
230        RUN;
231        %MACRO IDENT;
232        DATA NEWGEN;
233        MERGE IDEN GEN2;
234        RUN;
235        PROC SORT OUT=PHEN;
236        BY ID FAM;
237        RUN;
238        DATA PHEN;
239        MERGE PHEN FRAME;
240        PROC SORT OUT=PHEN;
241        BY POS;
242        RUN;
243        DATA PHEN;
244        MERGE ORIGIN PHEN;
245        MU=7.5;
246        HT=MU+GEN+HTRES;
247        RUN;
248        PROC CORR DATA=PHEN;
249        VAR GEN HT;
250        TITLE "CORRELATION FOR ITERATION &I AND DESIGN &DES H2=.2";
251        RUN;
252        PROC MEANS NOPRINT DATA=PHEN;
253        OUTPUT OUT=TEMP MEAN=HTMEAN;
254        VAR HT;
255        RUN;
256        DATA TEMP; SET TEMP;
257        DO T= 1 TO &TPOS;
258        HTMEAN=HTMEAN;
259        OUTPUT; END;
260        RUN;
261        DATA PHEN; MERGE PHEN TEMP;
262        KEEP REP FAM GEN GF GT HTRES HT HTMEAN POS ROW COL NAY;
263        RUN;
264        PROC PRINT DATA=PHEN;
265        TITLE "DESIGN &DES DATA SET";
266        RUN;
267        PROC SORT DATA=PHEN;
268        BY REP;
```

```
269        PROC MEANS NOPRINT DATA=PHEN;
270        OUTPUT OUT=TEMP MEAN=REPMEAN STD=REPSTD;
271        BY REP;
272        VAR HT;
273        RUN;
274        DATA TEMP; SET TEMP;
275        DO T=1 TO &NPOS;
276        REPMEAN=REPMEAN;
277        REPSTD=REPSTD;
278        STANDARD=REPMEAN+1.5*REPSTD;
279        OUTPUT;
280        END;
281        RUN;
282        DATA PHEN; MERGE PHEN TEMP;
283        FIRSTSEL=HT-STANDARD;
284        REPDEV=HT-REPMEAN;
285        KEEP REP FAM GEN GF GT HT HTMEAN NAY REPMEAN REPDEV STANDARD FIRSTSEL;
286        RUN;
287        PROC SORT DATA=PHEN OUT=PHEN;
288        BY FAM;
289        RUN;
290        PROC MEANS NOPRINT DATA=PHEN;
291        OUTPUT OUT=NCSEL MEAN=NCMN FAMMEAN;
292        BY FAM;
293        VAR FIRSTSEL HT;
294        RUN;
295        DATA NCSEL; SET NCSEL;
296        DO TRE=1 TO &NTREE;
297        NCMN=NCMN;
298        FAMMEAN=FAMMEAN;
299        OUTPUT; END;
300        RUN;
301        DATA PHEN; MERGE PHEN NCSEL;
302        RUN;
303        PROC SORT DATA=PHEN OUT=NCSEL;
304        BY DESCENDING NCMN DESCENDING FIRSTSEL;
305        RUN;
306        DATA NCSEL; SET NCSEL(OBS=&NSEL);
307        RUN;
308        DATA TEMP;
309        DO F=1 TO 10;
310        DO T=1 TO &NTREE; OUTPUT;
311        END; END; RUN;
312        DATA NCSEL; MERGE NCSEL TEMP;
313        RUN;
314        DATA NCSEL; SET NCSEL;
315        IF T<3 AND FIRSTSEL>0;
316        RUN;
317        PROC MEANS NOPRINT DATA=NCSEL;
318        OUTPUT OUT=NCSEL MEAN=NCSELMN GFMN GTMN;
319        VAR GEN GF GT;
320        RUN;
321        DATA COMB; MERGE NCSEL STANDARD;
322        PERCENT=(NCSELMN/BESTGEN)*100;
323        RUN;
324        PROC PRINT DATA=COMB;
325        TITLE 'SELECTIONS USING MODIFIED NC STATE METHOD';
326        RUN;
327        DATA GENCOR; SET PHEN;
328        IF TRE=1;
329        RUN;
330        PROC CORR;
331        VAR GF FAMMEAN;
332        TITLE 'CORRELATION BETWEEN FAMMEAN AND GF';
333        RUN;
334        PROC SORT DATA=PHEN OUT=PHEN;
335        BY NAY;
336        RUN;
337        PROC MEANS NOPRINT DATA=PHEN;
338        OUTPUT OUT=TEMP MEAN=NAYMN;
339        BY NAY;
```

```
340          VAR HT;
341          RUN;
342          DATA TEMP; SET TEMP;
343          IF NAY<113 THEN DO A=1 TO 25;
344          NAYMN=NAYMN;
345          OUTPUT; END;
346          ELSE IF NAY>112 THEN DO A=1 TO 20;
347          NAYMN=NAYMN;
348          OUTPUT; END;
349          RUN;
350          DATA PHEN; MERGE PHEN TEMP;
351          NAYDEV=HT-NAYMN;
352          RUN;
353          PROC SORT DATA=PHEN OUT=PHEN;
354          BY DESCENDING FAMMEAN;
355          RUN;
356          DATA PHEN; SET PHEN(OBS=&NSEL);
357          RUN;
358          %IF &DES=7 %THEN %GOTO T2;
359          %ELSE %IF &DES<7 %THEN %DO;
360          PROC SORT DATA=PHEN OUT=TEMP;
361          BY REP FAM;
362          RUN;
363          PROC MEANS NOPRINT DATA=TEMP;
364          OUTPUT OUT=TEMP MEAN=PLOTMEAN;
365          BY REP FAM;
366          VAR HT;
367          RUN;
368          DATA TEMP; SET TEMP;
369          P=%EVAL(&NTREE./&NREP.);
370          DO T=1 TO P;
371          OUTPUT; END;
372          RUN;
373          DATA PHEN; MERGE TEMP PHEN;
374          PLOTDEV=HT-PLOTMEAN;
375          RUN;
376          %END;
377          %T2:
378          PROC SORT DATA=PHEN;
379          BY FAM DESCENDING REPDEV;
380          RUN;
381          DATA MASSEL;
382          DO F=1 TO 10;
383          DO T=1 TO &NTREE;
384          OUTPUT; END;
385          END;
386          RUN;
387          DATA REPNAY; MERGE PHEN MASSEL;
388          IF T>10 THEN DELETE;
389          RUN;
390          PROC SORT DATA=REPNAY OUT=REPNAY;
391          BY FAM DESCENDING NAYDEV;
392          RUN;
393          DATA TEMP;
394          DO F=1 TO 10;
395          DO T=1 TO 10;
396          OUTPUT; END; END;
397          RUN;
398          DATA REPNAY; MERGE REPNAY TEMP;
399          IF T>2 THEN DELETE;
400          RUN;
401          PROC MEANS NOPRINT DATA=REPNAY;
402          OUTPUT OUT=REPNAY MEAN=COMBMN GFMN GTMN;
403          VAR GEN GF GT;
404          RUN;
405          DATA COMB; MERGE REPNAY STANDARD;
406          PERCENT=(COMBMN/BESTGEN)*100;
407          RUN;
408          PROC PRINT DATA=COMB;
409          TITLE 'MEAN GENETIC VALUE OF COMBINED REP AND NABORHOOD SELECTIONS';
410          RUN;
411          DATA MASSEL; MERGE PHEN MASSEL;
412          IF T>2 THEN DELETE;
```

```
413        RUN;
414        PROC MEANS NOPRINT DATA=MASSEL;
415        OUTPUT OUT=MASSEL MEAN=GENMASS GFMN GTMN;
416        VAR GEN GF GT;
417        RUN;
418        DATA COMB; MERGE MASSEL STANDARD;
419        PERCENT=(GENMASS/BESTGEN)*100;
420        RUN;
421        PROC PRINT DATA=COMB;
422        TITLE 'MEAN GENETIC VALUE OF REP DEVIATION SELECTIONS';
423        RUN;
424        DATA PHEN; SET PHEN;
425        PROC SORT DATA=PHEN OUT=PHEN;
426        BY FAM DESCENDING NAYDEV;
427        RUN;
428        DATA NAYSEL;
429        DO F=1 TO 10;
430        DO N=1 TO &NTREE;
431        OUTPUT; END;
432        END;
433        RUN;
434        DATA NAYSEL; MERGE NAYSEL PHEN;
435        IF N>2 THEN DELETE;
436        RUN;
437        PROC MEANS NOPRINT DATA=NAYSEL;
438        OUTPUT OUT=NAYSEL MEAN=NAYGEN GFMN GTMN;
439        VAR GEN GF GT;
440        RUN;
441        DATA COMB; MERGE NAYSEL STANDARD;
442        PERCENT=(NAYGEN/BESTGEN)*100;
443        RUN;
444        PROC PRINT DATA=COMB;
445        TITLE 'MEAN GENETIC VALUE OF NEIGHBORHOOD SELECTIONS';
446        RUN;
447        %IF  &DES=7 %THEN %GOTO T4;
448        %ELSE %IF &DES<7  %THEN %DO;
449        DATA PHEN; SET PHEN;
450        PROC SORT DATA=PHEN OUT=PHEN;
451        BY FAM DESCENDING PLOTDEV;
452        RUN;
453        DATA PLOTSEL;
454        DO F=1 TO 10;
455        DO W=1 TO &NTREE;
456        OUTPUT; END;
457        END;
458        RUN;
459        DATA PLOTSEL;
460        MERGE PLOTSEL PHEN;
461        IF W>2 THEN DELETE;
462        RUN;
463        PROC MEANS NOPRINT DATA=PLOTSEL;
464        OUTPUT OUT=PLOTSEL MEAN=PLOTGEN GFMN GTMN;
465        VAR GEN GF GT;
466        RUN;
467        DATA COMB; MERGE PLOTSEL STANDARD;
468        PERCENT=(PLOTGEN/BESTGEN)*100;
469        RUN;
470        PROC PRINT DATA=COMB;
471        TITLE 'MEAN GENETIC VALUE OF PLOT DEVIATION SELECTIONS';
472        RUN;
473        %END;
474        %T4:
475        %IF  &DES=3 OR &DES=6 OR &DES=7 %THEN %DO;
476        DATA NAYINDEX; SET PHEN;
477        NF=&NFAM;
478        H2=.4;
479        TN=.25;
480        NNAY=25;
481        CORR=.25;
482        TF=H2*CORR;
483        F=(1+TF*(NF-1))/NF;
484        H=(1+TN*(NNAY-1))/NNAY;
485        C=(1+TF*(NF-1)+CORR*(NNAY-1))/(NF*NNAY);
```

```
486        R=(1+CORR*(NF-1))/NF;
487        D=1/NNAY;
488        DET=F*H*(1+2*C-F-H)-C*C;
489        B1=H2*(F*H-C*C-R*H*(F-C)+F*D*(C-H))/DET;
490        B2=H2*(R*H*(1-H)-H*(F-C)+D*(F*H-C))/DET;
491        B3=H2*(F*(C-H)+R*(F*H-C)+D*F*(1-F))/DET;
492        INDEX=HT*B1+FAMMEAN*B2+NAYMN*B3;
493        RUN;
494        PROC SORT DATA=NAYINDEX OUT=NAYINDEX;
495        BY FAM DESCENDING INDEX;
496        RUN;
497        DATA INDEX;
498        DO F=1 TO 10;
499        DO T=1 TO &NTREE;
500        OUTPUT; END; END;
501        RUN;
502        DATA NAYINDEX; MERGE NAYINDEX INDEX;
503        IF T>2 THEN DELETE;
504        RUN;
505        PROC MEANS NOPRINT DATA=NAYINDEX;
506        OUTPUT OUT=NAYINDEX MEAN=INDSEL2 GFMN GTMN;
507        VAR GEN GF GT;
508        RUN;
509        DATA COMB; MERGE NAYINDEX STANDARD;
510        PERCENT=(INDSEL2/BESTGEN)*100;
511        RUN;
512        PROC PRINT DATA=COMB;
513        TITLE 'MEAN GENETIC VALUE OF SELECTIONS USING NEIGHBORHOOD INDEX';
514        RUN;
515        %END;
516        %MEND IDENT;
517        %IDENT
518        %END;
519        %MEND DESIGN;
520        %DESIGN(7, 1, 96, 2880)
521        %END;
522        %MEND GENET;
523        %MACRO BIG;
524        DATA ORIGIN;
525        INFILE HD;
526        INPUT COL 1-3 ROW 5-6 HTRES 8-14 4 NAY 16-18 POS 20-24;
527        %GENET(1, 1, 96, 1, 30, 2880)
528        %MEND BIG;
```

APPENDIX B

COMPUTER MODEL USED IN SELECTION

INDEX SIMULATIONS (SAS)

```
1        %MACRO NABOR(NRUN);
2        %LET TOTVAR=10;
3        %LET H2=.10;
4        %LET NTREE=45;
5        %LET NNAY=9;
6        %LET TN=.10;
7        %LET COR=.5;
8        %DO SIM=1 %TO 4;
9        %DO HERIT=1 %TO &NRUN;
10        %LET H2=&H2+.05;
11       %DO NAYCOR=1 %TO &NRUN;
12        %LET TN=&TN+.05;
13
14       DATA ONE;
15        SEED=0;
16        NF=&NTREE;
17        PHENVAR=&TOTVAR;
18        HRTBLY=&H2;
19        CORR=&COR;
20        FAMVAR=HRTBLY*PHENVAR;
21        FAMSTD=SQRT(FAMVAR);
22        MU=15;
23        DO FAM=1 TO 20;
24         GF=FAMSTD*RANNOR(SEED);
25         DO TREE=1 TO NF;
26          GT=FAMSTD*RANNOR(SEED);
27          GEN=GF+GT;
28          U=RANUNI(SEED);
29          OUTPUT; END;
30        END;
31       PROC SORT DATA=ONE OUT=THREE;
32        BY U;
33
34       DATA TWO;
35        SEED=0;
36        HRTBLY=&H2;
37        NAYSIZE=&NNAY;
38        NAYCOR=&TN;
39        PHENVAR=&TOTVAR;
40        ENVAR=PHENVAR-(HRTBLY*PHENVAR);
41        AMGVAR=NAYCOR*ENVAR;
42        AMGNAY=SQRT(AMGVAR);
43        WNAY=SQRT(ENVAR-AMGVAR);
44        DO NAYBER=1 TO 100;
45         NV=AMGNAY*RANNOR(SEED);
46         DO IND=1 TO NAYSIZE;
47          EN=WNAY*RANNOR(SEED);
48          OUTPUT; END;
49        END;
50
51       DATA FOUR; MERGE THREE TWO;
52        PHEN=MU+GF+GT+NV+EN;
53       PROC MEANS NOPRINT DATA=FOUR;
```

```
54          OUTPUT OUT=FIVE MEAN=NAYMN;
55          BY NAYBER; VAR PHEN;
56        DATA SIX; SET FIVE;
57          DO N=1 TO 9;
58          NHMN=NAYMN;
59          OUTPUT; END;
60        DATA SEVEN; MERGE FOUR SIX;
61        PROC SORT DATA=SEVEN OUT=EIGHT;
62          BY FAM;
63        PROC MEANS NOPRINT DATA=EIGHT;
64          OUTPUT OUT=NINE MEAN=FMN;
65          BY FAM; VAR PHEN;
66        DATA TEN; SET NINE;
67          DO T=1 TO 45;
68            FAMMN=FMN;
69            OUTPUT; END;
70
71        DATA ELEVEN; MERGE EIGHT TEN;
72          TF=HRTBLY*CORR;
73          F=(1+TF*(NF-1))/NF;
74          H=(1+NAYCOR*(NAYSIZE-1))/NAYSIZE;
75          C=(1+TF*(NF-1)+NAYCOR*(NAYSIZE-1))/(NF*NAYSIZE);
76          R=(1+CORR*(NF-1))/NF;
77          D=1/NAYSIZE;
78          C1=HRTBLY*(1-R)/(1-F);
79          C2=HRTBLY*(R-F)/(F*(1-F));
80          DET=F*H*(1+2*C-F-H)-C*C;
81          B1=HRTBLY*(F*H-C*C-R*H*(F-C)+F*D*(C-H))/DET;
82          B2=HRTBLY*(R*H*(1-H)-H*(F-C)+D*(F*H-C))/DET;
83          B3=HRTBLY*(F*(C-H)+R*(F*H-C)+D*F*(1-F))/DET;
84          INDEX1=PHEN*C1+FAMMN*C2;
85          INDEX2=PHEN*B1+FAMMN*B2+NHMN*B3;
86          DIFF1=(MU+GF+GT)-INDEX1;
87          DIFF2=(MU+GF+GT)-INDEX2;
88        PROC CORR DATA=ELEVEN;
89          VAR INDEX1 GEN;
90        PROC CORR DATA=ELEVEN;
91          VAR INDEX2 GEN;
92        PROC SORT DATA=ELEVEN OUT=TWELVE;
93          BY DESCENDING INDEX1;
94        PROC SORT DATA=ELEVEN OUT=THTEEN;
95          BY DESCENDING INDEX2;
96        DATA FRTEEN;
97          DO SEL=1 TO 45;
98          OUTPUT; END;
99        DATA FIFTEEN; MERGE TWELVE FRTEEN;
100         IF SEL<1 THEN DELETE;
101         KEEP HRTBLY FAM GEN NAYBER NV EN PHEN FAMMN NAYMN C1 C2
102              B1 B2 B3 INDEX1 INDEX2 DIFF1 DIFF2;
103       PROC PRINT DATA=FIFTEEN;
104         TITLE SELECTIONS USING INDEX1;
105       PROC MEANS NOPRINT DATA=FIFTEEN;
106         OUTPUT OUT=SIXTEEN MEAN=GENMN PHENMN;
107         VAR GEN PHEN;
108       DATA SVNTEEN; MERGE THTEEN FRTEEN;
109         IF SEL<1 THEN DELETE;
110        KEEP HRTBLY FAM GEN NAYBER NV EN   PHEN FAMMN NAYMN
111           C1 C2 B1 B2 B3 INDEX1 INDEX2 DIFF1 DIFF2;
112       PROC PRINT DATA=SVNTEEN;
113         TITLE SELECTIONS USING INDEX2;
114       PROC MEANS NOPRINT DATA=SVNTEEN;
115         OUTPUT OUT=EITEEN MEAN=GENMN PHENMN;
116         VAR GEN PHEN;
117       PROC PRINT DATA=SIXTEEN;
118         TITLE MEAN VALUE OF SELECTED TREES: FAMILY AND INDIV;
119       PROC PRINT DATA=EITEEN;
120         TITLE MEAN VALUE OF SELECTED TREES WITH NEIBORHOODS;
121       PROC MEANS DATA=FOUR;
122         OUTPUT OUT=NINETEEN MEAN=GENMN PHENMN;
123         VAR GEN PHEN;
124       PROC PRINT DATA=NINETEEN;
125         TITLE OVERALL MEANS FOR RUN=&HERIT.&NAYCOR;
```

```
126          %END;
127           %LET TN=.1;
128          %END;
129           %LET H2=.15;
130          %END;
131          %MEND NABOR;
131
```

l

VITA

Judith Loo-Dinkins

Candidate for the Degree of

Doctor of Philosophy

Thesis: COMPARISONS AMONG COMBINATIONS OF PROGENY TEST
DESIGNS AND SELECTION SYSTEMS FOR MAXIMUM GENETIC
GAIN IN LOBLOLLY PINE (PINUS TAEDA L.)

Major Field: Crop Science-Forest Resources

Minor Field: Statistics

Biographical:

Personal Data: Born on Prince Edward Island, Canada,
August 22, 1956, the daughter of Gerrit and Joyce
Loo; married Randy Dinkins, August 13, 1983;
Canadian citizen.

Education: Graduated from Kensington Rural High
School, Kensington, Prince Edward Island, June,
1974; received Bachelor of Science in Forestry
degree from the University of New Brunswick,
Fredericton, N.B., Canada, in May, 1979; received
Master of Science degree in Forestry from Oklahoma
State University, Stillwater, OK, in December, 1983;
and completed requirements for Doctor of Philosophy
degree from Oklahoma State University, in May, 1986.

Professional Experience: Forester with British
Columbia Ministry of Forests, 1979-1981; graduate
research assistant, Department of Forestry, Oklahoma
State University, 1981-1986.

Member: Canadian Institute of Forestry.