EFFECT OF PLOT SIZE ON ESTIMATES OF VARIANCE COMPONENTS IN WHITE PINE OPEN-POLLINATED HERITABILITY TESTS

William R. Gall and Eyvind Thor[1]

ABSTRACT

Two outplantings of an eastern white pine open-pollinated heritability test containing 113 and 128 families and replicated ten times, were used as uniformity trials to test the effect of plot size on the precision of several estimates. The standard errors of a plot mean, family mean and seed source mean on a per-tree basis were used as criteria of precision. The heritability estimate which contained the maximum amount of useful information was based on planting 7-tree family plots at the location which had 68% survival and on planting 9-tree plots at the location which had 33% survival. The seed source effect was estimated with the greatest precision at the high-survival location when 9-tree plots were used.

INTRODUCTION

Plot size and number of replications affect both the cost of an experiment and its capacity to provide a definitive answer to the questions being asked. Unfortunately, these two factors usually operate on an experiment in opposite directions, with one frequently suffering at the expense of the other. Thus, it is desirable to reach a compromise, to find the combination of plot size and number of replications which will yield the amount of information desired at the lowest cost.

The present study is analogous to a combination of a uniformity trial and a variety trial. The objective is to determine the plot size which will estimate the heritability of 5th-year total height of eastern white pine with the greatest useful precision.

PLANT MATERIALS

In establishing the experiment 10 mother trees of eastern white pine (*Pinus strobus* L.) were randomly selected from each of 13 southern Appalachian stands. Open-pollinated seed were collected, grown in the nursery for two years and outplanted in the spring of 1970 with 10 replications in each of four locations in Tennessee. Only two plantations, representing two physiographic regions of Tennessee, were used in this study. Ames Plantation, where one of the outplantings is located, is in the alluvial plain of the Mississippi River; Tullahoma, where the other outplanting is located, is in the Highland Rim region of Middle Tennessee. Ames Plantation is located 200 miles west of the

---

[1] Graduate research assistant and professor, respectively, Department of Forestry, The University of Tennessee, Knoxville, TN 37901

natural range of white pine, and survival there was poor: 33%. Tullahoma is just west of the southern Appalachian extension of the range of white pine, and survival was much better: 68%. When planted, there were 10 trees per plot but due to mortality the number of trees in any plot varied from 0 to 10. A total of 128 families were included in the Tullahoma plantation, while only 113 were planted at Ames.

In fall of 1974, following five growing seasons in the field, total height was measured, and this was the only variable analyzed.

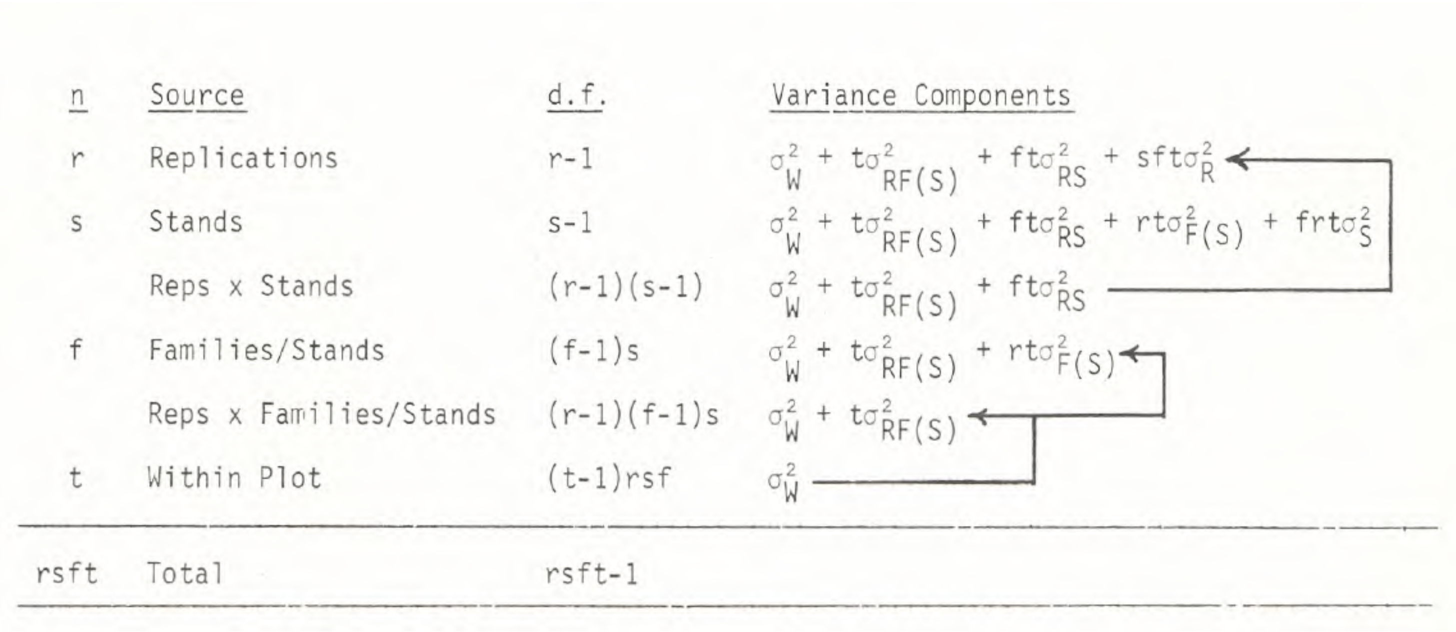
STATISTICAL METHODS

A FORTRAN program was written which would accept only a given number of trees per plot. Thus, in order to conduct an analysis on a data set which contained a maximum of n trees per plot, the original data set was fed into the FORTRAN program, which was set up for n trees per plot or less; the edited data set was written onto a disk and then read into the Statistical Analysis System (Service, 1972). Each location was analyzed individually. To create a data set which contained a maximum of n trees per plot, the first n living trees which appeared in each plot were selected. In the n-tree data set, after the first n trees in a plot were selected, all remaining trees in the plot were omitted from the analysis. In this method of selection, it is as though only the first n living trees in each plot are being measured. Thus, the 2-trees-per-plot data set was a subset of the 3-trees-per-plot data set, and so on. Consequently, every data set was correlated with every other data set, and the statistics derived from every data set must be correlated with those derived from every other data set. However, all data sets for one location should be independent of all data sets for the other location. Nine data sets were created and nine analyses were conducted for each location; no analysis was computed for the case of one tree per plot.

Analyses were carried out according to the model presented in Table 1, using Method I of Henderson, contained in the Statistical Analysis System. Henderson's Method I was used because, with such a large number of degrees of freedom, it would not have been possible to compute a least-squares analysis in one pass and, with 18 analyses, it would have been too time-consuming to partition the matrices so that portions of them could be inverted using the available memory capacity of the computer. No analysis produced negative sums of squares. As an additional labor-saving device, the arithmetic mean, rather than a harmonic mean, was used to compute the average number of trees per plot and the average number of families per stand for the purpose of solving for the variance components.

The heritability formula was the one used by Evans and Thor (1971). The variance components estimated by each mean square were derived by the method of Schultz (1955). F tests were performed as indicated in Table 1., except that a synthetic F test and degrees of freedom were computed by the method of Cochran (1951) to test the stand effect. The standard error of a plot mean, a family mean and a stand mean on a per-tree basis were obtained by dividing the appropriate error term by the average number of observations in each mean and taking the square root.

Table 1. --- Generalized analysis-of-variance table used in analyzing the Tullahoma and Ames white pine heritability test plantations.



n	Source	d.f.	Variance Components
r	Replications	r-1	$\sigma^2_W + t\sigma^2_{RF(S)} + ft\sigma^2_{RS} + sft\sigma^2_R$
s	Stands	s-1	$\sigma^2_W + t\sigma^2_{RF(S)} + ft\sigma^2_{RS} + rt\sigma^2_{F(S)} + frt\sigma^2_S$
	Reps x Stands	(r-1)(s-1)	$\sigma^2_W + t\sigma^2_{RF(S)} + ft\sigma^2_{RS}$
f	Families/Stands	(f-1)s	$\sigma^2_W + t\sigma^2_{RF(S)} + rt\sigma^2_{F(S)}$
	Reps x Families/Stands	(r-1)(f-1)s	$\sigma^2_W + t\sigma^2_{RF(S)}$
t	Within Plot	(t-1)rsf	σ^2_W
rsft	Total	rsft-1	

RESULTS AND DISCUSSION

All effects tested were significant at the 0.0005 level of probability for all plot sizes at the Tullahoma location. The replication, stand and rep x families/stand effects at the Ames location were also significant at the 0.0005 level of probability for all plot sizes, but the level of significance of the families/stand effect varied from a probability of 0.01 for 2-tree plots to a probability of 0.0005 for 10-tree plots.

Variance component and heritability estimates obtained for various plot sizes at the Tullahoma location are presented in Table 2; those for the Ames location are given in Table 3. All the variance components which contain environmental effects, that is, the within plot, the rep x families/stand and the replication estimates, are smaller for Tullahoma than for Ames. Correspondingly, variance components which contain primarily genetic effects, being the families/stand and stand estimates are larger for Tullahoma than for Ames. The Ames variance component estimates which contain environmental effects are larger even though the estimates of the overall mean were smaller at Ames than at Tullahoma.

The variance component estimates for the within-plot, families/stand and rep x families/stand effects, which are used in computing heritability estimates, are presented graphically in Figure 1. for each location along with the heritability estimates themselves. At each location, the within-plot estimate increases with increasing number of trees per plot, although it appears to increase faster at Tullahoma. The heritability estimates and the families/stand estimates remain very nearly constant with increasing plot size at both locations, although there may be a slight tendency for both to increase in the Ames analyses. The most striking differences between the two locations are the larger size of the within-plot and the rep x families/stand estimates at Ames and the larger size of the families/stand estimate at Tullahoma. These factors help to explain the higher heritability estimates at Tullahoma. None of the variance component estimates changes dramatically as plot size increases.

The standard error of a plot mean is presented in Figure 2; it decreases with increasing plot size at both locations. All plot sizes presented in these figures are maximum plot sizes; the corresponding average plot sizes (average number of living trees measured per plot) are given in Tables 2 and 3. These are both very smooth curves, probably due to the large number of plot means involved. The difference between the standard error of the mean of a 5-tree and a 6-tree plot in the Tullahoma analyses and that of a 4-tree and a 5-tree plot in the Ames analyses is less than a tenth of a foot. Since the trees were only measured to the nearest tenth of a foot, it would not be worthwhile to measure more than 5 trees in each plot in order to estimate the size of the rep x families/stand effect.

The standard error of a family mean is plotted against plot size in Figure 3. The difference between the largest standard error in the Ames analyses and the smallest one in the Tullahoma analyses is less than one tenth of a foot, so it would not be worthwhile to increase the plot size at either location above two trees per plot in order to estimate a family mean.

The same precision in estimating a family mean can be obtained in either location as long as total height is measured no closer than to one tenth of a foot. Due to the split-plot design and the large number of families, a family mean is estimated with the greatest precision which can be utilized.

Table 2. --- Variance component and heritability estimates obtained for various plot sizes in the Tullahoma plantation.

Maximum Plot Size	Average Plot Size	Variance Components						
		σ^2_W	$\sigma^2_{RF(S)}$	$\sigma^2_{F(S)}$	σ^2_{RS}	σ^2_S	σ^2_R	h^2
2	1.92	1.54	0.41	0.19	0.00	0.34	0.05	0.36
3	2.83	1.53	0.38	0.19	0.00	0.36	0.04	0.37
4	3.69	1.55	0.35	0.20	0.02	0.35	0.03	0.38
5	4.49	1.59	0.33	0.20	0.01	0.34	0.03	0.38
6	5.21	1.61	0.31	0.19	0.01	0.32	0.03	0.36
7	5.82	1.61	0.32	0.20	0.00	0.32	0.02	0.38
8	6.30	1.62	0.30	0.17	0.00	0.32	0.03	0.32
9	6.65	1.62	0.30	0.17	0.00	0.31	0.03	0.33
10	6.82	1.62	0.28	0.17	0.00	0.31	0.03	0.33

Table 3. --- Variance component and heritability estimates obtained for various plot sizes in the Ames plantation.

Maximum Plot Size	Average Plot Size	Variance Components						
		σ^2_W	$\sigma^2_{RF(S)}$	$\sigma^2_{F(S)}$	σ^2_{RS}	σ^2_S	σ^2_R	h^2
2	1.72	2.02	0.58	0.07	0.00	0.13	0.15	0.10
3	2.34	2.03	0.59	0.06	0.00	0.15	0.16	0.08
4	2.78	2.03	0.56	0.07	0.01	0.16	0.14	0.10
5	3.05	2.04	0.55	0.07	0.01	0.17	0.13	0.10
6	3.20	2.06	0.53	0.07	0.01	0.18	0.13	0.11
7	3.27	2.06	0.53	0.08	0.01	0.17	0.12	0.12
8	3.29	2.07	0.53	0.08	0.01	0.17	0.12	0.12
9	3.30	2.07	0.53	0.08	0.01	0.17	0.12	0.12
10	3.30	2.07	0.53	0.08	0.01	0.17	0.12	0.12

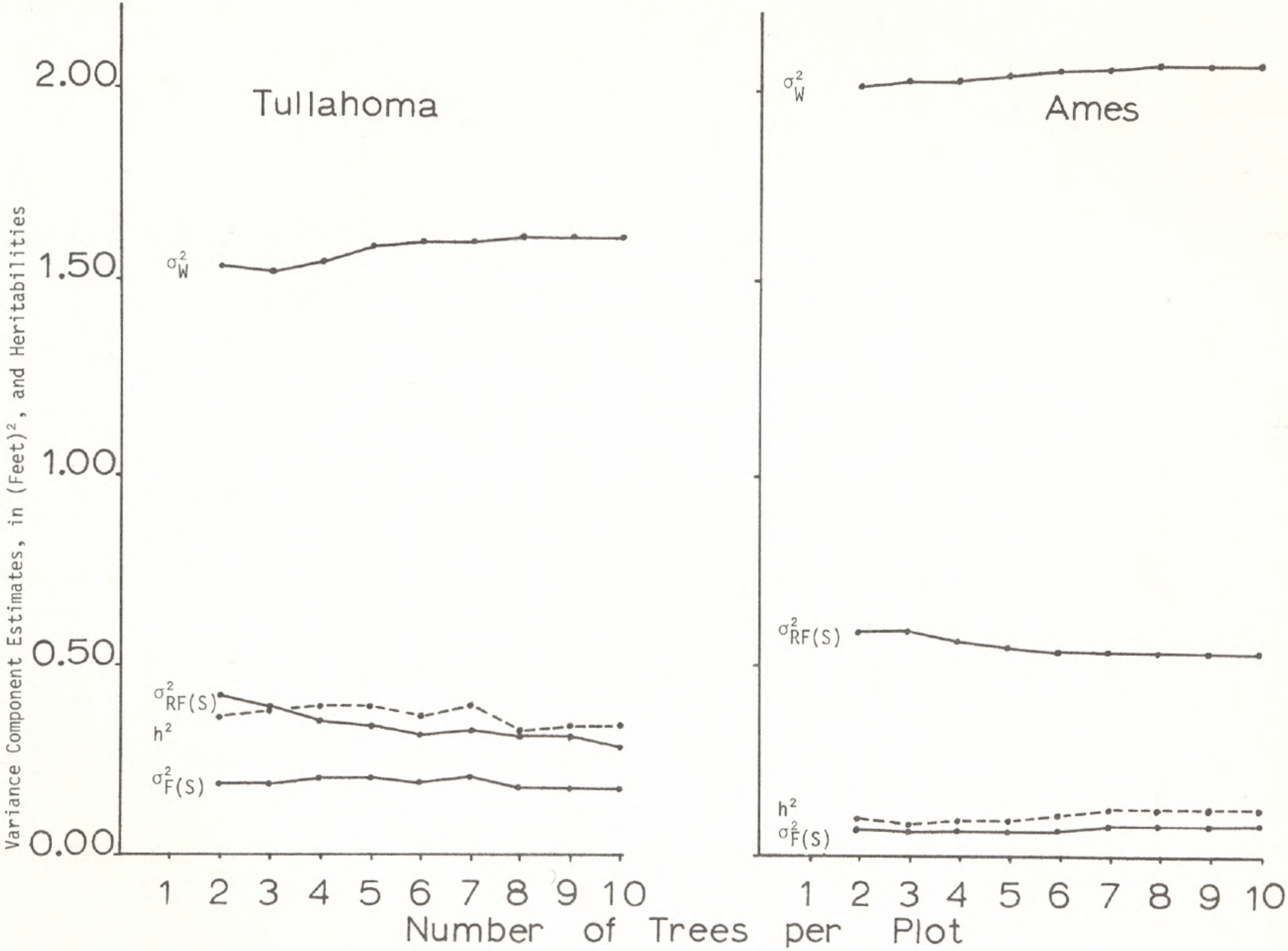


Figure 1. Estimates of heritabilities and variance components used to calculate heritabilities as a function of maximum plot size at the Tullahoma and Ames Plantation outplantings.

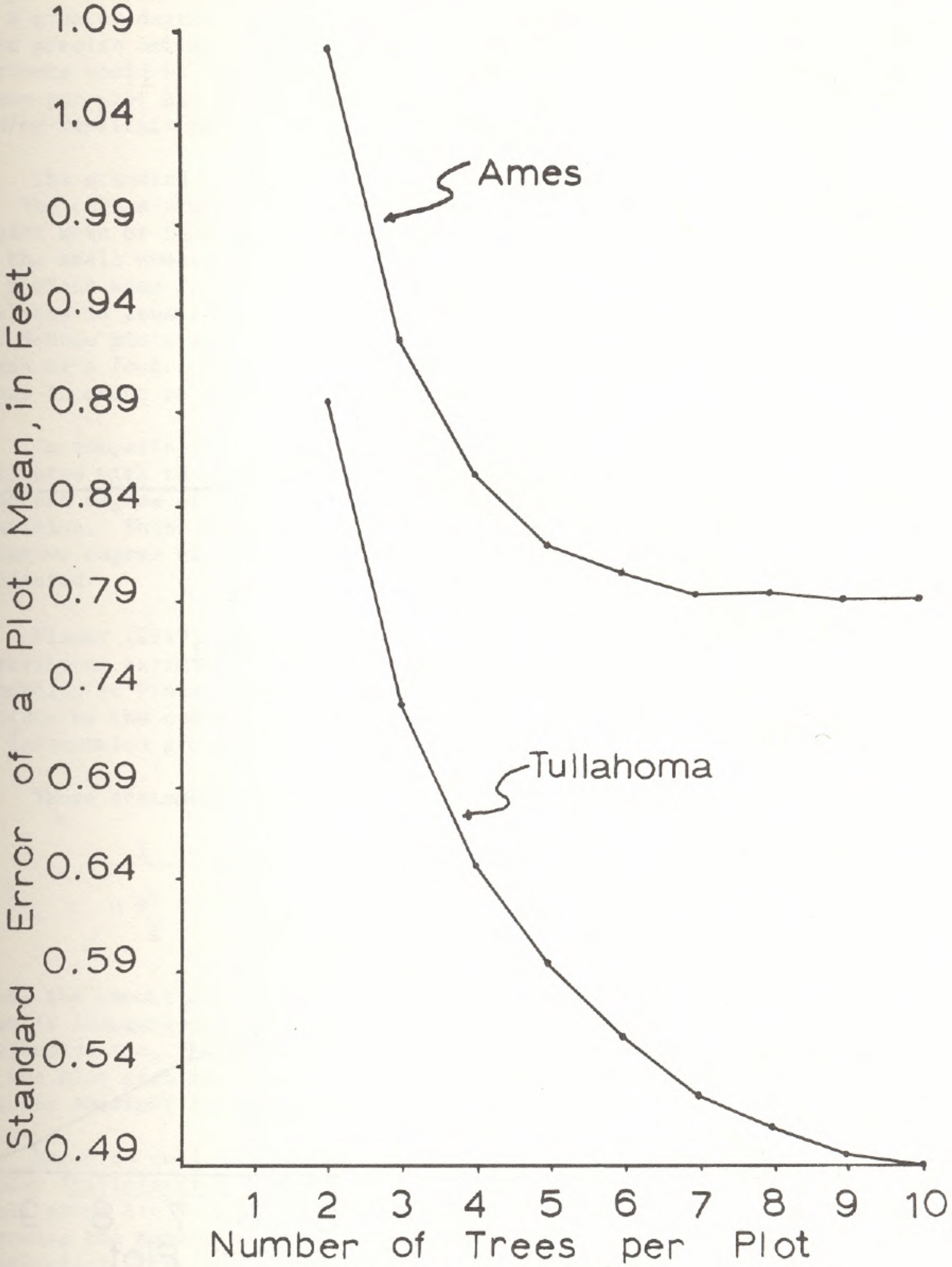


Figure 2. Estimates of the standard error of a plot mean as it varies with maximum plot size at the Tullahoma and Ames Plantation outplantings.

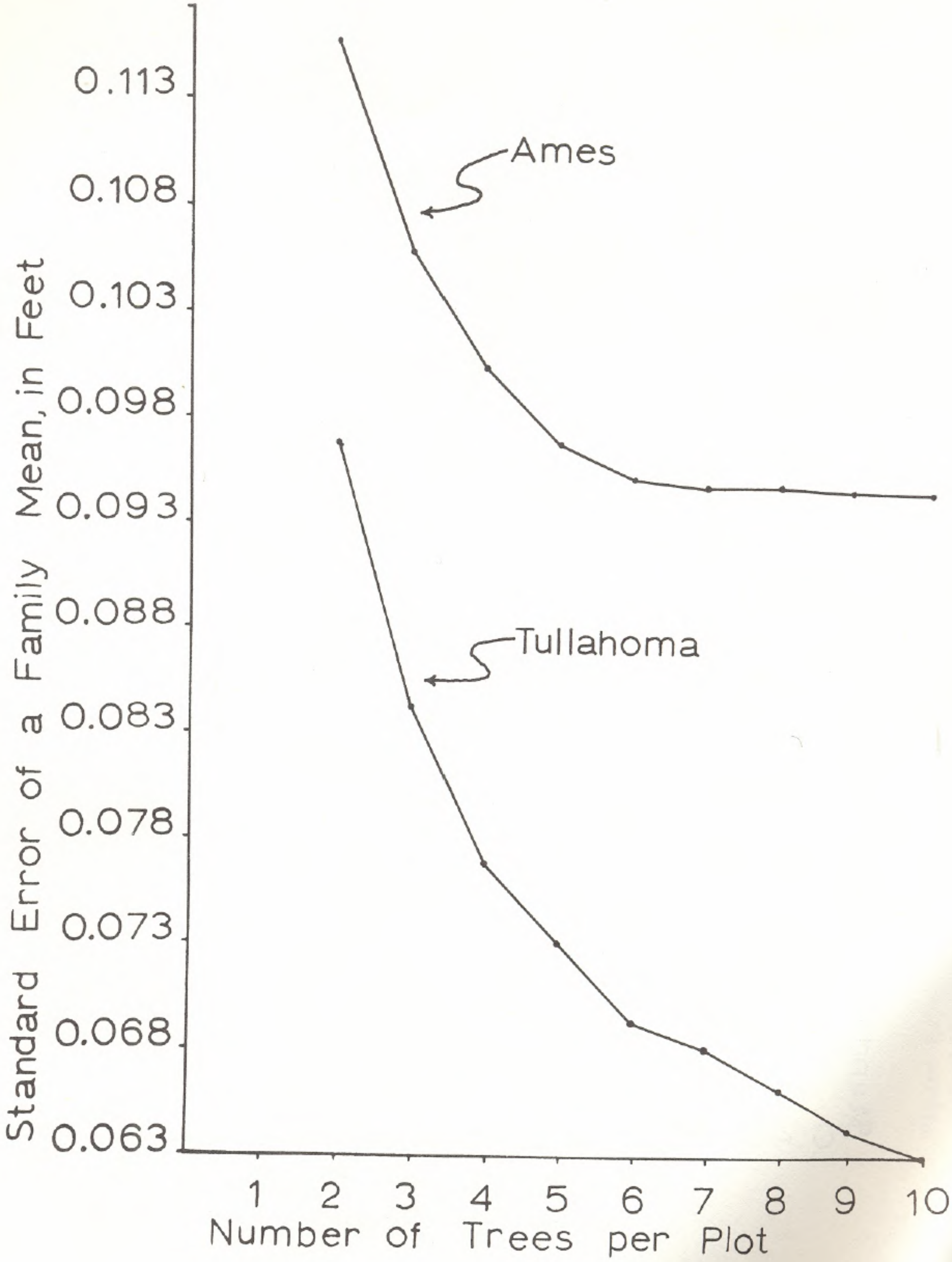


Figure 3. Estimates of the standard error of a family mean plotted against maximum plot size for the Tullahoma and Ames Plantation outplantings.

The standard error of a plot mean in the Tullahoma analyses is always smaller than that in the Ames analyses for the same size plot. Thus, the larger heritability estimates in the Tullahoma analyses are not the result of a greater degree of precision in estimating a family mean, but are due to more precise estimates of a plot mean. The precision of the heritability estimate could be increased by increasing the number of trees above two trees per plot but not by increasing the number of plots (additional families and/or replications).

The standard error of a stand mean is plotted against plot size in Figure 4. The curves are more irregular than those for either the standard error of a plot mean or that of a family mean. The irregularity can probably be explained by the small number of stand means. The difference between the standard error of a stand mean for 7-tree plots and that for 8-tree plots in the Tullahoma analyses is less than a tenth of a foot. The difference between the same quantity for 5-tree plots and that for 6-tree plots in the Ames analyses is less than a tenth of a foot. Thus, it would not be worthwhile to use plots larger than 7 trees in order to estimate the stand effect.

In comparing Figures 2, 3, and 4, it can be seen that a family mean is estimated with the greatest precision; a stand mean is estimated with the next greatest degree of precision, and a plot mean is estimated with the least precision. These results are in accord with theoretical evaluations of the relative degree of precision with which the main and split plot effects are estimated.

Fisher (1949) defines the amount of information provided by a normally distributed estimate as the reciprocal of the variance of the estimate. Again, according to Fisher, when the amount of information provided by an estimate is divided by the number of observations in the sample, it expresses the amount of information provided by each observation.

These statements may be expressed as:

$$\frac{1}{n\, s_{\bar{x}}^2} = \frac{1}{n\, \frac{s_e^2}{n}} = \frac{1}{s_e^2}$$

where s^2, is the sample variance, and n is the number of observations in each mean.

Thus, the amount of information per observation provided by an estimate of a mean is independent of the plot size if the sample variance remains constant. Obviously, too, the standard error of a mean will decrease as the square root of the plot size increases. This theoretical conclusion is borne out by the results obtained in this study.

Based on uniformity trials with red pine (Pinus resinosa Ait.) and yellow-poplar (Liriodendron tulipifera L.), Wright and Freeland (1959) concluded that small plots are more efficient, statistically, than large plots and that increasing the number of replications while using smaller plots increases the statistical efficiency if the same total number of trees is used.

Their plot sizes varied from one tree per plot to 64 trees per plot. Thus, even the maximum plot size possible in the present study is small compared to the range of plot sizes investigated by Wright and Freeland. Since the precision of the plct mean estimate was the primary determinant of the precision

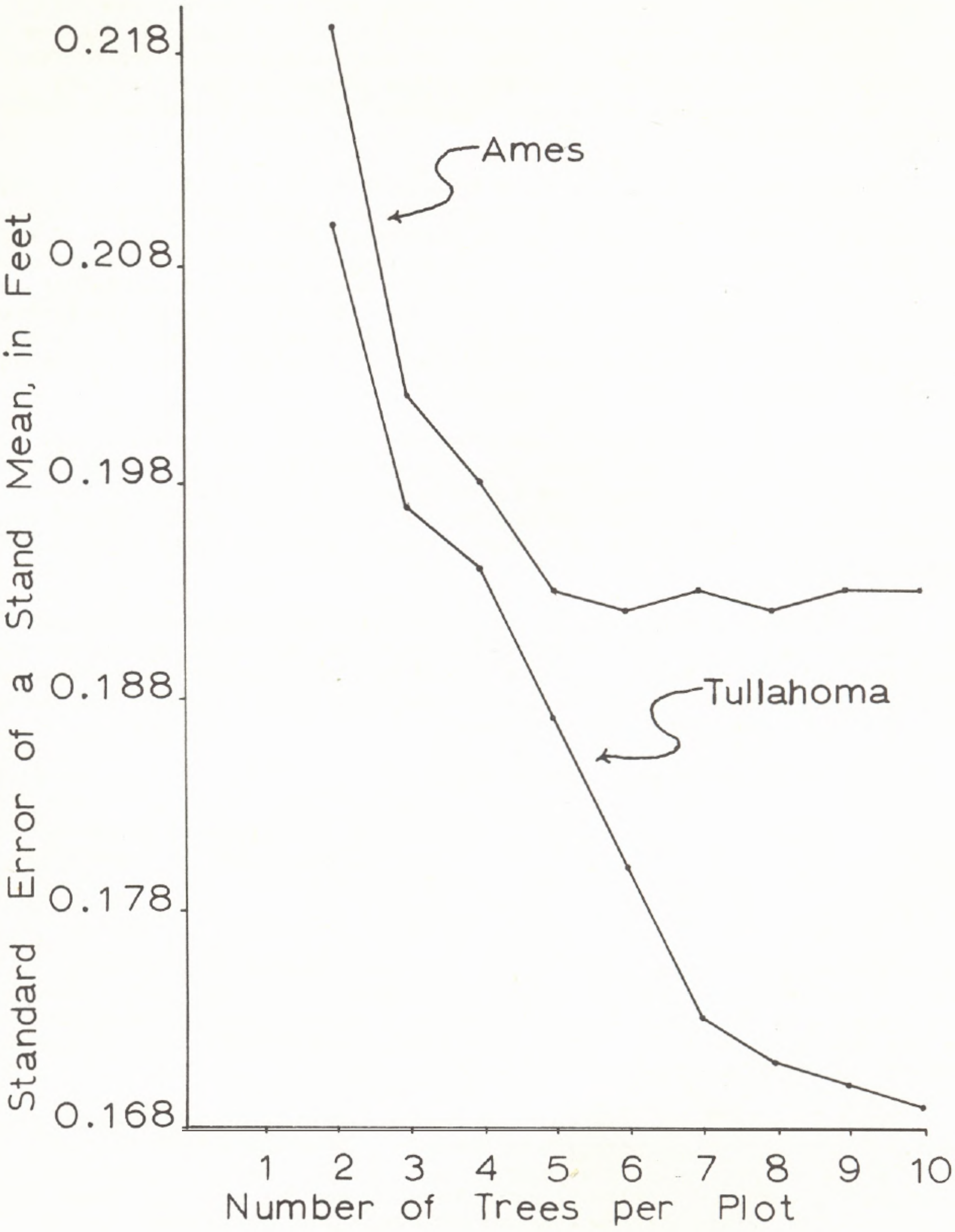


Figure 4. Estimates of the standard error of a stand mean versus maximum plot size for the Tullahoma and Ames Plantation outplantings.

of the heritability estimate in the present study, and since the maximum precision in estimating a plot mean was obtained when the maximum number of living trees measured was 5 trees per plot at Tullahoma and 4 trees per plot at Ames Plantation, the plot size which provided the most precise heritability estimate according to the present study was within the range of those recommended by Wright and Freeland.

Two differences between the two studies should be pointed out. First, Wright and Freeland were primarily concerned with determining optimum plot sizes for provenance experiments while the present study was conducted to assess the effect of plot size on heritability estimates. Secondly, the total number of trees in their experiments remained the same as plot size increased; in the present study, the total number of trees in the experiment increased each time plot size was increased, until all trees were included.

Conclusions

In Table 4. is presented the maximum number of living trees measured per plot in compiling each data set; in the first column under each location is shown the average number of living trees per plot which occurred in compiling each data set. In the second column under each location is given the number of trees per plot which would have to be planted in order to obtain the corresponding average number of living trees, based on the survival rate at each location. This table may be used to translate the results obtained in terms of maximum number of living trees measured per plot into the number of trees per plot which would have to be planted to attain the average plot size.

The greatest useful precision in estimating heritability was obtained at Tullahoma when a maximum plot size of 5 trees was used; thus, the most precise heritability estimate could have been achieved by planting only 7 trees per plot. The maximum amount of useful information needed to estimate heritability was produced at Ames Plantation when a maximum plot size of 4 living trees was used; thus, in order to obtain the maximum useful precision in estimating heritability, it would have been necessary to plant 9 trees per plot. It may be concluded that the planting of ten-tree plots in these plantations was reasonable for obtaining the most precise estimates of heritability for fifth year total height.

Table 4 --- Maximum and average plot sizes at Tullahoma and Ames Plantation with the number of trees necessary to plant to achieve the average plot size.

	Ames		Tullahoma	
maximum plot size	average plot size	plot size to plant	average plot size	plot size to plant
2	1.72	6	1.92	3
3	2.34	7	2.83	5
4	2.78	9	3.69	6
5	3.05	10	4.49	7
6	3.20	10	5.21	8
7	3.27	10	5.82	9
8	3.29	10	6.30	10
9	3.30	10	6.65	10
10	3.30	10	6.82	10

LITERATURE CITED

Cochran, W. G. 1951. Testing a linear relation among variances. Biometrics 7:17-32.

Evans, R. M. and E. Thor. 1971. Estimates of heritabilities and genetic gains for populations of Virginia pine. Proc. 11th Conf. on South. For. Tree Imp., Atlanta, Ga. p. 133-141.

Fisher, R. A. 1949. The Design of Experiments (5th ed.). Hafner Publishing Co., Inc., New York. p. 236.

Schultz, E. F., Jr. 1955. Rules of thumb for determining expectations of mean squares in analysis of variance. Biometrics 11:123-135.

Service, J. 1972. A User's Guide to the Statistical Analysis System. Student Supply Stores, N. C. State University, Raleigh, NC,260 pp.

Wright, Jonathan W. and F. Dean Freeland, Jr. 1959. Plot size in forest genetic research. Pap. of Mich. Acad. of Sci., Arts, and Letters 44: 177-182.