SAMPLING EFFICIENCY FOR 1-YEAR-OLD WHITE ASH SEEDLINGS

Chen Hui Lee and Calvin F. Beyl/

Abstract.--Measuring only the tallest seedling per plot was an acceptable sampling scheme for ranking family means in 1-year-old white ash. All of the deliberate schemes and all of the random schemes sampling 3 to 10 trees per plot were acceptable. Sample means were highly correlated with means computed by measuring every tree per plot (r values ranged from .93 to .98).

Additional keywords: random sampling, deliberate sampling, *Fraxinus americana* L., plot size, correlation coefficients.

In tree improvement work, it is often necessary to measure the same trees year after year. This can involve a lot of time and money. We suggest that partial measurement of plots may be adequate for some types of research.

Partial measurement of nursery genetic data on ponderosa pine (*Pinus ponderosa* Laws) suggests that measuring the tallest seedling per plot was adequate and that a deliberate sampling scheme was more efficient than a random sampling scheme (Lee 1974). To test the validity of partial measurement on another tree species, we collected data from a white ash (*Fraxinus americana* L.) experiment at the Illinois Division of Forestry Tree Nursery at Jonesboro, Illinois. In contrast to the ponderosa pine study, the 1-0 white ash trees were more variable in height and in the number of trees per plot (density).

For partial measurement analyses we selected seedlings in two families from each of 16 natural stands. The stands were located in an area extending from 30.3° N (Texas) to 45.7° N (Wisconsin) and from 68.9° W (Maine) to 94.4° W (Texas).

A randomized complete block design with three replications was used in the nursery. Each plot consisted of seedlings in a plot 30 cm x 120 cm. Number of seedlings per plot ranged from 11 to 107.

MATERIAL AND METHODS

At the end of the first growing season, the height of each seedling was measured to the nearest 3 cm and plot means were computed in 21 different ways, i.e., based on: all seedlings; the tallest and shortest per plot (T + S); the 1, 2, 3, 4, 5, 6, 8 or 10 tallest seedlings in each plot; the tallest tree from each of 10-, 5-, or 3-tree groupings (conveniently referred to as 10, 20, or 30 percent sampling, respectively) per plot; and 1, 2, 3, 4, 5, 6, 8 or 10 seedlings randomly chosen with replacement in each plot.

1/ Associate Professor of Forestry, College of Natural Resources, University of Wisconsin-Stevens Point, Wisconsin 54481, and Principal Plant Geneticist, North Central Forest Experiment Station, Forestry Sciences Laboratory, USDA Forest Service, Carbondale, Illinois 62901.

RESULTS AND DISCUSSION

Plot means based on all trees ranged from 12 to 75 cm with individual tree heights from 15 to 150 cm. Within the same plots, many of the tallest trees were three times as tall as the shortest.

ANALYSIS OF VARIANCE

To test for differences among stands and families/stands, height data were analyzed following a randomized complete block design with the hierarchal arrangement. Means based on all seedlings in each plot were used in the analysis. The corresponding degrees of freedom and variance components were obtained as follows:

Source of variation	D.F.	MS	F	Expected mean square
Stand (S)	s-1=15	M_1	M_1/M_2	$\sigma^2+n\sigma^2_{F:S}+fn\sigma^2_S$
Families within stand (F:S)	s(f-1)=16	M_2	M_2/M_3	$\sigma^2+n\sigma^2_{F:S}$
Trees within family (T:F)	fs(n-1)=64	M_3		σ^2

There were significant differences among stands but not among families /stands (table 1). The lack of significance for families might be due to the limited number (2) of families per stand. The test points out the importance of geographic variation in white ash tree improvement.

The ANOVA was also run for all the sampling schemes. There were significant differences among stands for all the deliberate schemes and for half of the random schemes. There were no significant differences among families within stands for any of the sampling schemes.

VARIANCE COMPONENTS

To understand what happens with the various sampling schemes, we computed the means and variance components for each. We expected the stand and F:S component to stay about the same with the various schemes, while the T:F component should decrease with higher sampling intensity. In the deliberate sampling scheme, the stand component stayed about the same for the various sampling intensities. There was greater variation in stand variance using random sampling, suggesting that random sampling of only a few trees may give erratic estimates. In deliberate sampling the T:F component decreased as the number of trees sampled increased. In random sampling the T:F component was about the same for plots of 3 to 10 trees, but higher for 1- and 2-tree plots.

The variation in seedling height of white ash *was* much greater than for ponderosa pine. Using data from all trees per plot, we got a coefficient of variation of 21% for white ash compared with 3% for ponderosa pine. White ash height growth differences among and within plots were noticeable even to a casual observer, whereas this was not the case with ponderosa pine.

Table 1.--Average height, F-value, and variance components for 21 sampling schemes

Sampling scheme		Average height (cm)	F values stands[a]	Variance components: Stand	F:S	T:F
Deliberate	T+S	38	4.05	9.0	0	58.2
	1	68	4.30	30.9	0	189.5
	2	63	3.73	30.7	0	177.7
	3	62	4.06	31.5	0	168.6
	4	59	4.36	34.7	0	153.7
	5	58	4.21	32.1	0	150.0
	6	56	4.33	33.3	0	145.7
	8	53	4.07	31.7	0	133.4
	10	51	3.87	30.9	0	125.6
	10%	53	3.86	26.1	0	149.1
	20%	47	4.19	21.9	0	115.0
	30%	42	3.80	21.2	0	90.1
Random	1	30	5.03	27.6	0	113.4
	2	33	2.06	11.4	0	92.8
	3	31	2.43	8.9	0	64.4
	4	32	1.86	8.3	0	62.2
	5	31	1.80	9.8	3.4	64.0
	6	32	3.01	16.4	0	60.2
	8	32	2.06	9.6	0	57.9
	10	32	3.42	13.4	0	60.0
All trees		32	2.95	13.2	0	46.7

[a] F values of 2.35 or larger are needed for significance at .05 level.

CORRELATION ANALYSIS

To determine the amount of genetic information that might be lost because of partial measurements, we ran a correlation analysis. Simple correlations were computed between family means based on measurement of all trees and family means based on measurements of only a sample (table 2). Deliberate sampling had a higher correlation than random sampling for 1- and 2-tree plot schemes, but not for 3 or more trees per plot. Under deliberate sampling, we got essentially as much information measuring the tallest tree as measuring all trees, and of course much more economically. The coefficient of determination (r^2) for all vs. 1 tree/plot was .91. In other words, 91% of the information is still available by measuring only about 2% of the trees. There was no additional gain by measuring 2 to 10 trees per plot and very little gain by measuring the tallest and the shortest or 10, 20, or 30%. If random sampling is practiced, one should sample at least three trees. There is a 28% loss of information when sampling only one tree and 23% loss when sampling only two trees per plot. At 3 to 10 trees per plot, the information loss is reduced to about 10%. For 3 to 10 trees it doesn't matter whether random or deliberate sampling is used.

Table 2.--Correlation coefficients for various plot sizes and sampling schemes

Sampling scheme & tree per plot	r	Sampling scheme & tree per plot	r
All vs Ran 1	0.85	All vs Delib 1	0.95
All vs 2	0.88	All vs 2	0.96
All vs 3	0.95	All vs 3	0.95
All vs 4	0.95	All vs 4	0.96
All vs 5	0.95	All vs 5	0.96
All vs 6	0.97	All vs 6	0.95
All vs 8	0.93	All vs 8	0.95
All vs 10	0.96	All vs 10	0.94
		All vs T + S	0.97
		All vs 10%	0.97
		All vs 20%	0.98
		All vs 30%	0.97

The 5 percent point of r with 14 degrees of freedom = 0.50.

The 1 percent point of r with 14 degrees of freedom = 0.62.

If the principles derived here are applicable to outplantings, deliberate sampling may save measurement time with only a minor loss of information. It is usually easy to locate the tallest tree in each plot. Although it has not been tested, we expect that the principle would also hold for other quantitative traits. The reader should be aware of two other aspects of deliberate sampling: (1) it does not provide an absolute average value for each plot or genetic unit, and (2) it is necessary to make a correction if heritability is to be computed (Kung 1976).

In comparing white ash with ponderosa pine, we see that deliberate sampling of the tallest tree/plot will generally be acceptable for both species, even though white ash is so much more variable than ponderosa pine. Apparently the amount of variation does not greatly affect the efficiency of a sampling scheme. Further tests using outplanting data with white ash and other species will be needed before we can make sampling recommendations beyond our nursery results.

LITERATURE CITED

Kung, Fan H. 1976. Adjustment and interpretation of progeny test when only the best tree in each plot is measured. Paper for Western Forest Geneticists Association, Centralis, Washington, Aug. 5, 1976.

Lee, C. H. 1974. Effect of partial measurement on statistical precision and efficiency in a nursery study of Pinus ponderosa. Silvae Genet. 23:124-126.