

MEASUREMENT OF PARTIAL REPLICATIONS AND ITS
EFFECT ON THE ASSESSMENT OF GENETIC DATA

By Chen Hui Lee, Professor of Forestry
College of Natural Resources, University of Wisconsin
Stevens Point, Wisconsin 54481

ABSTRACT--Random numbers were used to create height, dbh and wood specific gravities data assumed to have come from a simulated permanent study plantation. They were subjected to analysis following a randomized complete block design with 30 half-sib families in each of 10 replications. An analysis of variance was performed separately by varying the number of replications (two through ten) using the random and deliberate sampling schemes. For each analysis, the variance components for family, blocks and error were obtained and family heritabilities estimated.

The inclusion of only two replications in the sample was found inadequate; there should be at least six replications sampled or available in the study plantation for the assessment of genetic information. They may be chosen either randomly or deliberately. The justification for the reduced replication approach was discussed.

INTRODUCTION

Working with 30 Lake States jack pine (Pinus banksiana Lamb.) seed source material, I failed to obtain any geographic variation pattern in the selected growth and wood properties. The study material was furnished by the Institute of Forest Genetics, North Central Forest Experiment Station, US Forest Service, Rhinelander, Wisconsin. It came from Plantation No. 12 located in Pembine, northern Wisconsin and was 22 years old from seed in 1978. Those data were generated from one tree representing each of 30 seed sources in each of two replications. The inclusion of inadequate number of replications in the sample was suspected of causing that problem.

In the absence of the additional study material, it was felt necessary to find an alternative means for further exploration. I turned to the use of the random numbers and created three sets of data simulating the jack pine range in height, dbh and wood specific gravities. This paper reports the outcome of a computer simulation.

MATERIAL AND METHODS

The "IRANDOM" program available through the Minitab Student Handbook (Duxbury Press, North Scituate, Massachusetts by Ryan, Joiner and Ryan) was used to create three sets of data on height, dbh and wood specific gravities (Appendices I through III, available for distribution upon request). Those data were analyzed according to the randomized complete block design with 30 half-sib families in each of 10 replications. There are $f-1$, $n-1$ and $(f-1)(n-1)$ degrees of freedom for family, blocks and error respectively. Plot (seedlot) means were used as items in the analysis.

For each trait, an analysis of variance was performed separately by varying the number of replications from two to ten chosen randomly without replacement or deliberately (rep. 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9 and 1-10). The analysis was done according to the following tabulation:

Source of variation	D.f.	MS	F	Expected MS
Family	$f - 1$	M_1	M_1/M_3	$\sigma_e^2 + n\sigma_F^2$
Blocks	$n - 1$	M_2		$\sigma_e^2 + f\sigma_R^2$
Error	$(f-1)(n-1)$	M_3		σ_e^2

The purpose of the correlation analysis was two-fold: to determine the amount of loss in genetic information and to check whether the varied number of replications sampled affects the ranks of family means. There are 28 degrees of freedom and family means based on measurement of varied number of replications were used as items in the analysis.

RESULTS AND DISCUSSION

Analysis of variance-- Partial measurement of the genetic material established either in nursery or permanent plantation may be done for convenience after the cost-and-benefit factor being carefully analyzed. It can be achieved by the reduction of the plot size (measuring one or two best trees in each plot) or by the reduction of the number of replications (sampling portion of the nursery or plantation material) or both. The latter approach has received little studies and attention.

The results of this computer simulation showed a distinct trend that the sampling of the study material from only two replications selected either randomly or deliberately was not adequate to serve

the study purpose of forest geneticist and tree breeders (Table 1). It was true regardless of what trait was studied. If we accept the face value of the statistical analysis, we would be misled to conclude that a given tree species lacks the genetic variation. Failure to recognize the genetic diversities in tree species bears undesirable consequences: we may overlook some valuable breeding material and the risk may be greater when working with non-local or exotic species whose silviculture is relatively unfamiliar to us.

When I increased the number of replications in the working sample, it began to demonstrate the existence of the significant between-family differences in the three simulated growth and world properties. It would appear that at least six replications would be needed for the purpose of the assessment of genetic information. It did not matter whether they were selected at random or deliberately.

It is interesting to know how the family, blocks and error variance components change as the number of replications was gradually increased in the analyses. The computation of family heritabilities (h^2) shown in Table 2 was conducted according to the following formula:

$$h^2 = \sigma_F^2 / (\sigma_R^2/n + \sigma_F^2)$$

The impact of the varied sampling intensities (combination of sampling scheme and replication number) is obvious: there was a general improvement in the estimate of family heritabilities as more replications were poured into the sample. It may be necessary to adjust the size of the family heritabilities whenever partial measurement is employed in the analysis (Kung, 1977).

Standard error of a plot (seedlot) mean was useful to define the sampling efficiency in my present study. The similar technique was used in a white pine (Pinus strobus L.) study in Tennessee (Gall, et al., 1977). I calculated the standard error by dividing the error variances by the number of replications (n), then square rooted them. The results are presented in Figure 1.

As expected, the size of the standard error tended to decrease as more replications were sampled. When the standard error were plotted against the number of replications, the plateau of the curve occurred when six or more replications were included in the sample. This means the size of the standard error remained stabilized. That trend persisted regardless of what trait was studied and what sampling scheme was used in the analysis.

Correlation analysis-- Simple correlations were computed between family means based on measurement of all 10 replications and family means based on measurement of reduced replications selected

Table 1. Test of significance for height, dbh and wood specific gravities as affected by the sampling scheme and number of replications in the sample

Sampling scheme	No. of replications	F values		
		Height	dbh	Sp. gr.
Random	2	0.641	1.186	1.590
	3	2.092**	1.545	1.369
	4	1.393	1.403	0.971
	5	1.429	1.279	0.933
	6	1.752*	1.390	1.849**
	7	1.493	1.734*	1.647*
	8	2.389**	1.809**	1.467
	9	2.237**	1.748**	2.013**
	Deliberate	2	1.481	0.798
3		0.996	1.174	0.915
4		0.888	0.987	1.001
5		1.043	1.176	1.218
6		1.262	1.111	1.569*
7		1.624*	1.439	1.668*
8		1.751*	1.543*	2.063**
9		2.014**	1.467	2.455**
All ten			2.386**	1.634*

* Significant at the 5 percent level

** Significant at the 1 percent level

Table 2. Estimate of family heritabilities as affected by varied number of replications and different sampling schemes

Sampling scheme	No. of replications	Family heritabilities		
		Height	dbh	Sp. gr.
Random	2	0	0.15	0.39
	3	0.52	0.35	0.28
	4	0.29	0.29	0
	5	0.31	0.23	0
	6	0.47	0.29	0.45
	7	0.33	0.42	0.42
	8	0.57	0.43	0.35
	9	0.55	0.42	0.50
	Deliberate	2	0.32	0
3		0	0.15	0
4		0	0	0
5		0.05	0.17	0.13
6		0.19	0.11	0.37
7		0.39	0.30	0.39
8		0.41	0.35	0.54
9		0.50	0.31	0.61
All ten			0.58	0.38

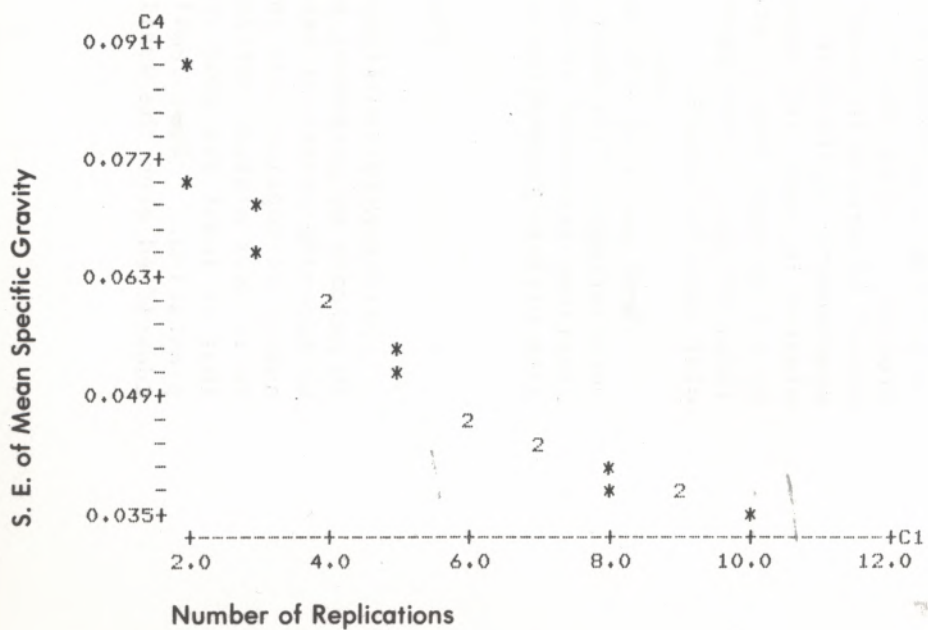
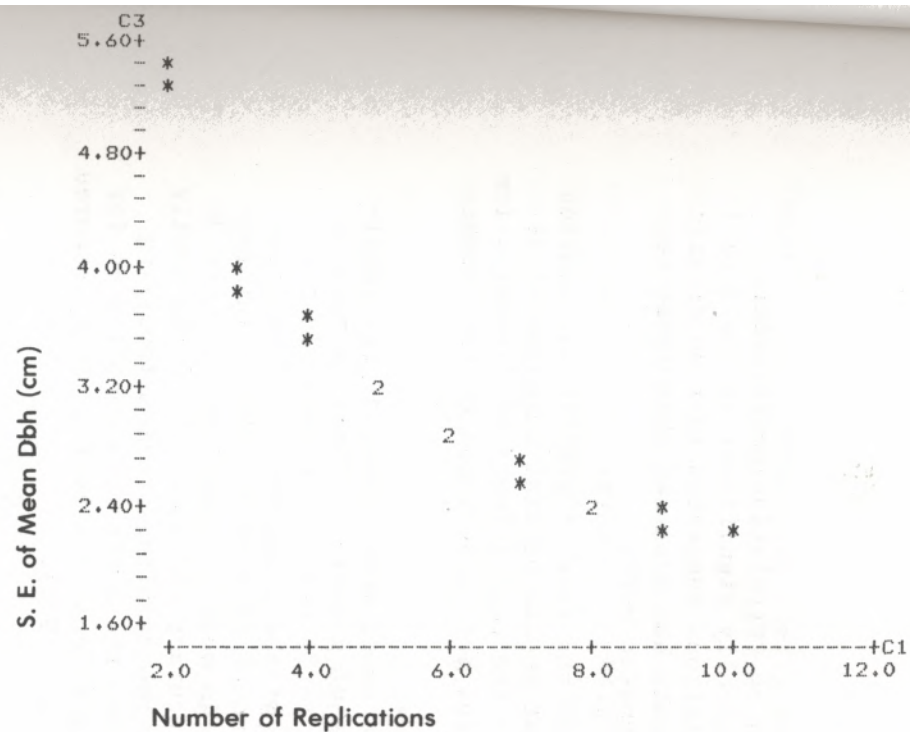
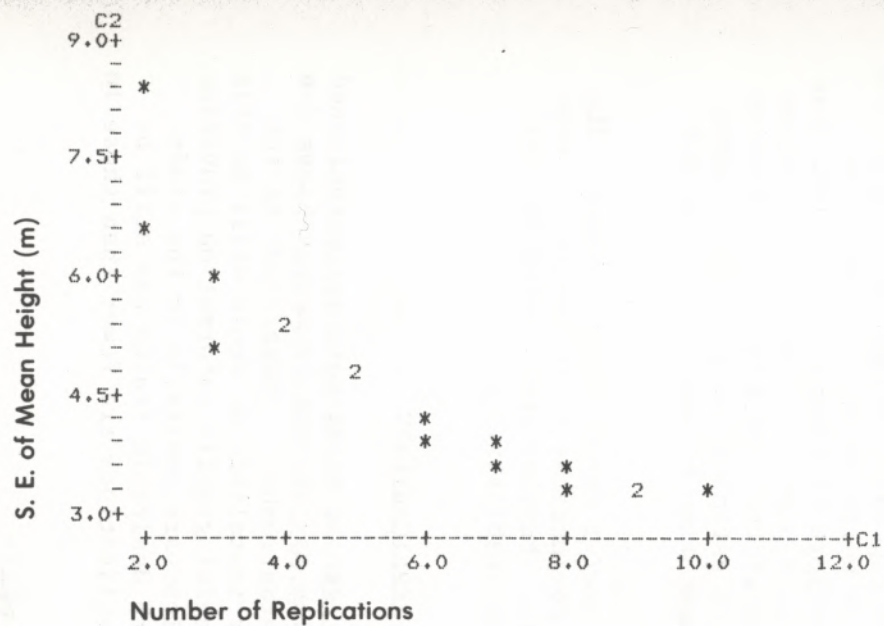


Figure 1. The magnitude of the standard error of plot mean as varied with the sampling schemes and the number of replications sampled.

randomly or deliberately. Of the 48 correlation coefficients shown in Table 3, 46 were statistically significant at the 5 or 1 percent level. The strong correlations suggested that no drastic changes in the ranks of family means had occurred when fewer replications were sampled for the genetic work.

It was necessary to determine the size of genetic information on family means that might be lost due to partial measurement of replications. The subtraction of the coefficients of determination (r^2) from one was considered useful for measurement of the amount of loss in genetic information.

When the random sampling was used, measurement of six replications was accompanied by a loss of genetic information ranging from 21 percent in height growth to 31 percent in dbh. In the deliberate sampling, measurement of the equal number of replications resulted in a comparable amount of loss in genetic information: it ranged from 24 percent in wood specific gravities to 30 percent in height growth. The accuracy of estimate for the family means did not differ between the two sampling schemes involving six or more replications. However, the deliberate sampling is far superior over the random sampling if only the best tree is measured in each plot (Lee, 1974; Lee et al., 1976).

Measurement of six replications yielded a saving of 40 percent in measurement effort while it was accompanied by a loss of 20 to 30 percent in genetic information depending on what trait was studied. This practice is still worth trying, however, the amount of saving in measurement effort is substantially lower than measurement of the best tree in each plot. If there are 12 trees planted in each plot, measurement of the best tree is accompanied-- by a 4 percent loss in accuracy of estimate for the family means (Lee, 1974). Saving in measurement time becomes greater as the plot size increases.

Both sampling schemes were useful for my study purpose. The correlation coefficients tended to increase as the number of replications increased in the sample. Further tests using factual data will be needed for wider applicability.

PRACTICAL APPLICABILITY

Frequently, portion of the genetic plant material established in nursery or permanent plantation may be lost or made useless due to man-made causes or natural catastrophe. Regardless of the number of replications initially installed, we would still be able to recover a great portion of useful genetic information providing that at least six good replications are available in the study plantation. Best families with desirable traits can still be identified even the available replications are fewer than those in

Table 3. The correlation coefficients comparing the effect of different replications sampled by two sampling schemes

Sampling scheme	No. replications	Height growth	dbh	Wood specific gravity
Random	10 vs 2	0.678	0.542	0.609
	3	0.782	0.564	0.725
	4	0.809	0.626	0.678
	5	0.763	0.803	0.764
	6	0.887	0.832	0.839
	7	0.916	0.871	0.922
	8	0.940	0.918	0.970
	9	0.976	0.962	0.975
	Deliberate	10 vs 2	0.303	0.631
3		0.414	0.612	0.571
4		0.625	0.691	0.712
5		0.755	0.825	0.801
6		0.838	0.860	0.874
7		0.898	0.884	0.919
8		0.952	0.918	0.948
9		0.980	0.970	0.974

The 5 percent point of r with 28 degrees of freedom = 0.361
 The 1 percent point of r with 28 degrees of freedom = 0.463

the original establishment.

LITERATURE CITED

- Gall, W. R. and E. Thor
1977 EFFECT OF PLOT SIZE ON ESTIMATES OF VARIANCE COMPONENTS
IN WHITE PINE OPEN-POLINATED HERITABILITY TEST, PLANTATIONS.
24th Northeast. For. Tree Improve. Conf. Proc.
pp 29-40.
- Kung, F. H.
1977 ADJUSTMENT AND INTERPRETATION OF PROGENY TESTS WHEN
ONLY THE BEST TREE IN EACH PLOT IS MEASURED. *Silvae
Genet.* 26: 117-119.
- Lee, C. H.
1974 EFFECT OF PARTIAL MEASUREMENT ON STATISTICAL PRECISION
AND EFFICIENCY IN A NURSERY STUDY OF PONDEROSA PINE.
Silvae Genet. 23: 124-126.
- Lee, C. H. and C. F. Bey
1976 SAMPLING EFFICIENCY FOR 1-YEAR-OLD WHITE ASH SEEDLINGS.
10th Central States For. Tree Improve. Conf. Proc.
pp 105-108.