COMPARISONS OF SAMPLING METHODS FOR INVENTORY OF NURSERY STOCK

R. E. Mullin
Ontario Department of Lands and Forests
Maple, Ontario

This study was conducted to meet the demand of our nurserymen for a comparison of the relatively simple and cheap procedures of systematic sampling with those of random and stratified random sampling which were considered in a previous report of this Department (12). In the last two methods of sampling, sample size, number, and sampling technique were considered. The present samplings were concerned with only the number of living trees. In this discussion the term "systematic" indicates that the first count was taken at a random placement and all subsequent counts at equal intervals. It is not claimed that this procedure is the equivalent of random sampling (9, 10).

In much of the literature it is implied that a systematic sample may be more precise than either of the other types (2 5, 8, 10, 13, 15). Precise is used here to refer to the size of deviations from the mean, m, and accurate is used to refer to deviations from the true mean, u. Warning is frequently given that a fully valid estimate of the precision of a systematic sample as indicated by error (as standard error of the mean, confidence limits, or error percentage) cannot be readily calculated (2, 5, 6, 7, 15, 18).

Several mathematical devices have been suggested for obtaining an estimate of the error in systematic sampling (2, 3, 4, 10, 18). Some authors have suggested the use of supplementary random counts to provide information on the sampling error (10, 15).

Shiue (13) and Shiue and John (14) presented formulae for estimation of error, based on multiple random starts of systematic sampling, relating this type of sampling to "cluster" sampling as a specialized type of stratified sampling. However, such a procedure is complicated in field practice and computation. There may still be some hesitancy about accepting its validity.

Finney (5) noted that it was a rather poor procedure to analyze data from a systematic sample in the same manner as a random sample. Nevertheless, Yates (18) and Grosenbaugh (8) observed that if this procedure is followed, an overestimate of the sampling error will usually be obtained.

The following advantages of systematic sampling have been indicated by various authors:

1. It is easier to apply in the field, saves time, and can be done by less-skilled workers.
2. Because it is simpler, less mistakes may be made.
3. The sampling units are widely and evenly distributed and should be representative of all parts of the population.

The disadvantages generally given are:

1. The sample may give a poor estimate of the mean if a periodicity exists in the population which is in tune with spacing of the samples.
2. There is no fully valid method of computing error.

Sampling Recommendations for Nurseries

Johnston (11) studied sampling for tree nursery inventory based on total counts of several seed and transplant beds. He presented comparisons of systematic with random and stratified random samples. He concluded that for broadcast conifer seedbeds the smallest size sampling unit tested, 0.5 feet by bedwidth, was most satisfactory. Similarly, for conifer transplants he found that the 1-foot bed-width sample (also the smallest tested) was the most efficient. In reference to sampling method he states that for "seedbed stock the systematic samples proved to be more accurate than stratified random samples, but in the transplant beds neither method was consistently more accurate than the other."

Wakeley (17), in references to nurseries for southern pines, recommended the use of "across the bed" frames (1 foot by 4 feet) for sampling seedbeds. He stated "All sample locations must be drawn strictly at random, with absolutely no exercise of personal judgement," thus excluding all systematic sampling.

Stoeckeler and Jones (16) gave tables of the numbers of counts per bed required to obtain a specified percent of error. Their recommendations implied a stratified random system for seedbeds. However, where five or more counts are required per bed, they suggested pacing at equal intervals. They did not specify random placement of the first counts, and did not refer to this as systematic sampling. For transplants they gave tables for selecting the number of 6-foot row counts in a stratified random procedure.

Barton and Clements (1) described a sequential systematic sampling technique for nursery inventories. After taking a fixed number of counts, 20, at fixed intervals, they applied a formula to compute the number of additional counts required to bring the estimate to within ± 10 percent. They noted "There is no recognized statistical method for testing the degree of accuracy of the systematic sampling discussed in this article." They did not compare the method with random or stratified random procedures but observed that it has been found reliable in 6 years of practical use.

Methods

In the earlier report of this Department (12), recommendations were presented for an inventory of seven categories of nursery stock in Ontario nurseries. This report was based on procedures of random and a stratified random sampling and included sample computations of standard deviation, standard error of the mean, and error percentage. (The last term is defined in the computations which follow.) The report contained recommendations for size, number of counts, and method of sampling for estimations at two levels, within 5 percent for shippable stock and 10 percent for other stock.

The present study is based on the comparison of the above procedures with systematic sampling, as applied to the most abundant category in our nurseries, machine transplants. For this category the earlier manual suggested 80 counts with a 4-foot frame (4 feet of bed length across the 6 or 7 rows of transplants) or 120 counts with a 2-foot frame in a stratified random pattern, both to give an error percentage of less than five. The bed, which may be 500 to 1,500 feet long, is the unit of stratification.

In this study, 15 nursery units (blocks of one species and age class), with an approximate range of 150,000 to 900,000 trees, were sampled in each of the following six ways

1. Fully random, 4-foot frame. Approxi mately 80 counts.
2. Fully random, 2-foot frame. Approxi mately 120 counts.
3. Stratified random, 4-foot frame. Approx imately 80 counts.
4. Stratified random, 2-foot frame. Approx imately 120 counts.
5. Systematic (random start), 4-foot frame. Approximately 80 counts.
6. Systematic (random start), 2-foot frame. Approximately 120 counts.

A schematic drawing of the methods is given in figure 1.

Although 80 and 120 were used as a guide, the number of 'counts varied somewhat. In the fully random and systematic 4-foot samplings, 80 counts were made in 22 of the 30 examples, and in the other examples, 78 to 82 counts were taken. In the fully random and systematic 2-foot samplings, 120 counts were made in 24 of the 30 examples, and 116 to 121 counts were taken in the remaining six examples. In the stratified random samplings, restricted by a fixed number per bed, the variation was greater; in the 4-foot samples it was 70 to 100, and in the 2-foot samples it was 100 to 144.

To reduce the dangers of bias, a crew of two reliable students did all the counting and measuring. Rigid sampling frames and steel measuring tapes helped make the counting exact and eliminated bias in selection of frame location.

Results

The nurseryman is chiefly interested in the mean of his counts. As a guide to the reliability of his mean, he can measure the precision of

SAMPLING METHODS

MACHINE TRANSPLANTS

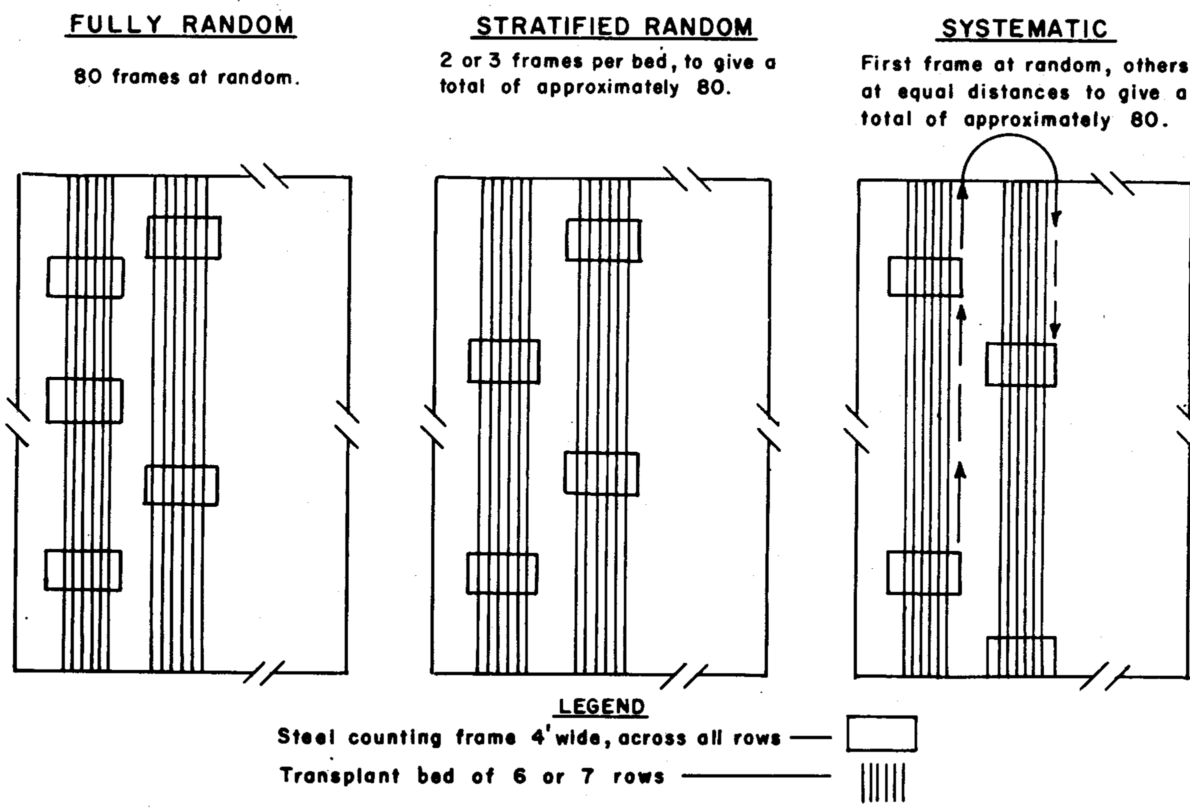

Figure 1.--Six methods of sampling that were used.

the sample, i.e., the spread of the counts around -their mean, m. This can be done by means of the standard error or by error percentage (defined later).

In the section on the comparison by error percentage (below) of the two counting frames and three sampling methods, it is assumed that the nurseryman, to be practical, should specify a fixed number of counts to his counting crews. Thus, the standard error, adjusted by a correction when counts greater or less than 80 for 4-foot frames or 120 for 2-foot frames, and the error percentage are used. In a later comparison, an alternative form of computation, calculation of the number of counts required to meet a specified error percentage, . is presented.

Comparison by error percentage

The results of analysis of machine trans

plants from 15 nursery units are summarized in table 1. The fully random and stratified random samplings have been computed in the manner shown in the following examples.

Example 1. Computation for fully random.

Nursery unit: Canal Field.

No. of samples (y), N = 80
Total count(sigma)y = 7,447

Sample mean = 93.0875

SS or sigma(y^2) = 706,177

TABLE 1.--Results of 6 samplings of machine transplants of 15 nursery units[1]

Nursery unit	4-foot frames						2-foot frames						General mean[2]
	Fully random		Stratified random		Systematic		Fully random		Stratified random		Systematic		
	Mean	Error percent	Mean	Error percent	Mean	Error percent	Mean (x2)	Error percent	Mean (x2)	Error percent	Mean (x2)	Error percent	
H-22.......	105.45	2.12	100.83	3.41	100:54	2.30	102.98	3.24	104.10	1.34	105.08	2.35	103.23
H-29.......	103.22	2.75	102.80	1.73	103.12	1.61	101.74	2.60	103.40	1.74	102.05	1.49	102.66
Canal......	93.09	3.06	95.01	3.21	92.92	3.32	96.85	3.16	95.00	2.34	93.97	3.16	94.64
T-33.......	89.85	3.94	90.86	3.08	87.84	3.35	89.97	3.47	92.12	3.26	94.37	2.43	91.13
H-21.......	105.95	2.64	105.04	2.87	103.36	2.89	106.05	2.45	104.89	2.79	106.58	2.46	105.40
T-72.......	71.12	5.16	68.26	4.88	68.15	5.41	66.23	5.44	69.61	5.36	70.68	5.25	68.96
T-12.......	94.43	3.26	90.51	3.21	91.73	2.59	95.10	3.19	92.75	2.29	96.32	2.25	93.56
H-13.......	94.52	3.74	87.37	3.57	90.40	4.24	89.83	3.89	89.90	2.92	89.32	3.78	90.14
H-16.......	105.32	2.67	104.92	1.86	104.85	3.04	107.67	2.69	108.20	1.24	108.58	1.66	106.90
H-15.......	80.64	3.93	78.26	4.16	77.94	4.93	79.90	3.63	82.87	3.45	83.42	2.88	80.85
H-12.......	91.60	3.57	97.75	2.42	98.09	2.49	98.72	2.77	98.37	2.57	97.77	2.99	97.28
T-14.......	90.96	3.84	88.22	2.44	87.92	3.89	89.52	3.79	89.63	2.45	90.08	2.66	89.48
Bl. 1......	108.22	1.66	108.61	1.39	108.61	1.45	108.85	2.15	109.83	1.30	108.97	1.95	108.90
Bl. 2......	112.85	1.28	113.26	1.38	110.62	1.24	111.67	1.47	112.54	1.55	114.47	1.60	112.66
Bl. 3......	114.16	2.46	120.50	1.25	116.51	1.15	119.44	1.74	120.61	1.38	117.69	1.47	118.37
Averages	97.425	3.07	96.813	2.72	96.173	2.93	97.637	3.04	98.255	2.40	98.623	2.56	97.611

[1] Comparisons were by means and error percentages.
[2] General mean = total frame counts for all methods divided by total number of counts.

$$\text{Variation} = \Sigma(y^2) - \frac{(y)^2}{N} = 12{,}954.39$$

$$\text{Variance} = \frac{\text{variation}}{N-1} = \frac{12{,}954.39}{79} = 163.980$$

$$\text{Standard deviation} = \sqrt{163.980} = 12.80546$$

$$\text{Standard error} = \frac{\sigma}{\sqrt{N}} = \frac{12.80546}{8.94427} = 1.432$$

(N of 80 or 120 was used in all cases in computing standard error).

$$\text{Confidence limit} = \text{Standard error} \times t = 1.432 \times 1.99 = 2.849$$

$$\text{Error percent} = \frac{2.849}{93.0875} = 3.06 \text{ percent}$$

Example 2. Computation for stratified random.

Nursery unit: Canal Field.

No. of samples (y), N = 84

Two per bed in each of 42 beds.

Total count $\Sigma y = 7{,}981$

Sample mean = 95.0119

SS or $\Sigma(y^2) = 771{,}671$

Sum of squares for bed totals (two samples per bed) = 1,527,531

$$\text{Total variation} = 771{,}671 - \frac{63{,}696{,}361}{84} = 13{,}380.99$$

$$\text{Within bed variation (for 42 d.f.)} = 771{,}671 - \frac{1{,}527{,}531}{2} = 7{,}905.5$$

$$\text{Between bed variation (for 41 d.f.)} = 13{,}380.99 - 7{,}905.5 = 5{,}475.49$$

$$\text{Variance} = \frac{7{,}905.5}{42} = 188.226$$

$$\text{Standard deviation} = \sqrt{188.226} = 13.71955$$

$$\text{Standard error} = \frac{\text{standard deviation}}{\sqrt{N}} = \frac{13.71955}{8.94427} = 1.534$$

(N of 80 or 120 was used in all cases in computing standard error).

$$\text{Confidence limits} = \text{standard error} \times t = 1.534 \times 1.99 = 3.053$$

$$\text{Error percent} = \frac{3.053}{95.0119} = 3.21 \text{ percent}$$

Example 3. Computation for systematic.

The counts were grouped in 2's and 3's as in the stratified random sampling, and the same form of computation was used. The counts were entered in order on the tally sheets, thus permitting this form of stratification. As noted before, this method of computation is not fully valid.

Comparison by number of counts

The same computation as used in the comparison by error percentage was used to obtain the standard error. The standard error was used to compute the number of counts required to obtain a specified error percentage, using the formula N = 400n error percent. The results are shown in table 2. Again, this comparison incorporates the weakness of nonvalidity but only for the systematic samples.

TABLE 2.--Results of 6 samplings of 15 nursery units of machine transplants[1]

Nursery unit	4-foot frames			2-foot frames		
	Fully random	Stratified random	Systematic	Fully random	Stratified random	Systematic
H-22	14.4	37.2	16.9	50.4	9.3	26.5
H-29	24.2	9.6	8.3	32.5	14.5	10.6
Canal	30.0	33.0	35.3	47.9	26.3	47.9
T-33	49.7	30.4	35.9	57.8	51.0	28.3
H-21	22.3	26.4	26.7	28.8	37.4	29.0
T-72	85.2	76.2	93.7	42.0	137.9	132.3
T-12	34.0	33.0	21.5	48.8	25.2	24.3
H-13	44.8	40.8	57.5	72.6	40.9	68.6
H-16	22.8	11.1	29.6	34.7	7.4	13.2
H-15	49.4	55.4	77.8	63.2	57.1	39.8
H-12	40.8	18.7	19.8	36.8	31.7	42.9
T-14	47.2	19.1	48.4	68.9	28.8	34.0
Bl. 1	8.8	6.2	6.7	22.2	8.1	18.2
Bl. 2	5.2	6.1	4.9	10.4	11.5	12.3
Bl. 3	19.4	5.6	4.1	14.5	9.1	10.4

[1] Comparisons were by number of counts required for 5 percent error.

Test of equality of the variances

A test of the equality of the variances was made among the three combinations of fully random/stratified random, fully random/systematic, and stratified random/systematic, within the two sizes of sampling frame. About half of these tests showed significance, some to the 0.01 percent level.

Size of sampling frame

In each of the three sampling techniques, the average number of trees per frame for all 15 nursery units is greater for the 2-foot frame (doubled to make it comparable) than for the 4-foot frame (table 1). In an analysis of variance of the means (fully valid as not based on standard error), the differences are significant at almost the 0.1 percent level, and the difference between the 4-foot and 2-foot frames accounts for about 70 percent of the variation.

A second comparison, although less reliable because of inclusion of standard error, was made in terms of the error percentages (table 1). The average error percentage is less for the 2-foot frames than for the 4-foot frames (significant at almost the 1.0 percent level). Also, the number of counts of 2-foot frames required to reach the same percentage was less than twice the number of 4-foot frames (table 2). This implies that the 2-foot frame was more precise.

A third comparison was made in terms of the "general mean" of table 1. This was obtained by totaling the approximately 600 counts in each nursery unit, with 2-foot frames doubled, and by dividing by the number of counts. Lacking a total count, the general mean was taken as the closest available estimate of the true mean. This does not entirely preclude bias. It was found that in 11 of the 15 nursery units the means which were farthest from the general mean (underlined in table 1) were taken by 4-foot sampling frames.

The above considerations imply that the 2-foot frame is more satisfactory.

The most likely explanation is that there is less tendency to miss seedlings in the smaller frames. The problem of "border effects" (bias for including or excluding trees on the side margins of the frames), which would be expected to be double for smaller frames in relation to numbers of trees counted, appears to be negligible.

Method of sampling

The data in table 1 represent six separate attempts to obtain an estimate of the number of trees in each nursery unit. In an analysis of variance, the average number of trees per frame failed to show significance for the differences due to sampling method. There was no definite bias for one method to give higher or lower averages, and hence estimates of the populations.

However, when the error percentages were analyzed, the differences due to sampling methods were significant at greater than the 0.1 percent level. The overall average error percentages were: Fully random, 3.06 percent; stratified random, 2.56 percent; and systematic, 2.74 percent. The significance was due largely to the difference between the fully random and the other two methods, which did not differ significantly from each other. The implication is that the fully random is the least reliable system and the stratified the most reliable, with the systematic being intermediary, when they are judged according to error terms. This implication is supported by the information in table 2.

One further comparison is possible. The means of each sampling method can be compared with a mean of all counts. Because of differences between the 4-foot and 2-foot frame, this comparison has to be made separately fo the two frame sizes. In the 30 comparisons it was found that the sampling systems yielding means farthest from the general mean were as follows: Fully random, 14; stratified random, 7; and systematic, 9.

The foregoing indicates that the stratified random sampling of this nursery category is best, particularly if a valid error is required. However, from the practical standpoint, it can easily be seen from table 1 that the means obtained by systematic sampling are, on the average, within 1 percent of those obtained by the stratified sampling. In rare instances this difference may be 3 or 4 percent, but the nurseryman may accept this as reasonable, particularly as he may have a safety factor of 5 or 10 percent which he applies to his estimate before listing on the inventory.

Literature Cited

1. Barton, William Warren, and Clements, Charles M.
   1961. A systematic sampling nursery inventory procedure. Tree Planters' Notes 46: 19-25.
2. Cochran, William G.
   1953. Sampling techniques. 330 pp. New York: John Wiley & Sons Inc.
3.
   1956. In Statistical methods, by George W. Snedecor, pp. 489-523. Ames, Iowa: Iowa State College Press.
4. DeLury, D. B.
   1950. Values and integrals of the orthogonal polynomials up to n = 26. Univ. of Toronto Press, 33 pp.
5. Finney, D, J.
   1948. Random and systematic sampling in timber surveys. Forestry 12(1): 64-99.
6. Finney, D. J.
   1953. An introduction to statistical science in agriculture. 179 pp. New York: John Wiley & Sons Inc.
7. Fisher, R. A.
   1953. The design,of experiments. Ed. 6, 244 pp. London: Oliver and Boyd.
8. Grosenbaugh, L. R.
   Forest parameters and their statistical estimation. Paper presented at Auburn Polytechnic Inst. Conf. on Statistics Applied to Res., Auburn, Ala., Sept. 8,1948.
9. Huntsberger, David V.
   1961. Elements of statistical inference, 291 pp. Boston: Allyn and Bacon Inc.
10. Jeffers, J. N. R.
    1960. Experimental design and analysis in forest research, 172 pp. Stockholm: Almquist and Wiksell.
11. Johnson, Floyd Alfred.
    1943. A statistical study of sampling methods for tree nursery inventories. Jour. Forestry 41: 674-679.
12. Mullin, R. E., Morrison, L. M., and Schweitzer, T. T.
    1955. Inventory of nursery stock. Ontario Dept. of Lands and Forests. Res. Rpt. 33, 64 pp.
13. Shiue, Cherng-Jiann..
    1960. Systematic sampling with multiple random starts. Forest Sci. 6(1): 42-50.
14. ____________ and John, Hugo H.
    1962. A proposed sampling design for extensive forest inventory. Double systematic sampling for regression with multiple random starts. Jour. Forestry 60: 607-610.
15. Spurr, Stephen H.
    1952. Forest inventory, 476 pp. New York: The Ronald Press Company.
16. Stoeckeler, J. H., and Jones, G. W.
    1957. Forest nursery practice in the Lake States. Forest Service, U. S. D. A., Agr. Handbook 110, 124 pp.
17. Wakeley, Philip C.
    1954. Planting the southern pines. Forest Service, U. S. D. A., Agr. Monograph 18, 233 pp.
18. Yates, Frank.
    1960. Sampling methods for censuses and surveys, 3d Ed., 440 pp. London: Charles Griffin & Co. Ltd.