Estimating Merchantable Seedlings in Nursery Seedbeds

Cuauhtemoc Saenz-Romero, Trent L. Marty, and Raymond P. Guries

Professor, Universidad Michoacana de San Nicolks de Hidalgo, Instituto de Investigaciones sobre los Recursos Naturales, Morelia, Michoacan, Mexico, Forest Tree Improvement Specialist, Wisconsin Department of Natural Resources, Bureau of Forestry, Madison, WI, and Professor, University of Wisconsin-Madison, Department of Forest Ecology and Management, Madison, WI

We determined the sample size needed to accurately estimate the number of merchantable seedlings per seedbed in forest tree nurseries in Wisconsin. Analysis of highly variable seedling stocking within some seedbeds indicates that it is better to use more, small, closely spaced samples distributed throughout a seedbed, rather than a few, large, widely spaced samples distributed across a seedbed. We suggest a new sampling design using 4-ft-long (1.22-m-long) row segments distributed throughout the seedbed, where the number of samples (row segments) can be adjusted proportionally to the heterogeneity of seedling density within the seedbed. Our approach is broadly applicable to other bareroot nurseries. Tree Planters' Notes 50(1): 23-27; 2003.

Estimating the number of merchantable seedlings prior to lifting is an important activity for any bareroot forest tree nursery. An accurate inventory is needed to balance the available nursery stock with customer orders. If the inventory is underestimated, customer orders must be canceled; if it is overestimated, surplus stock must be maintained for another year or be destroyed. In either instance, additional costs are incurred.

Current inventory procedures at State forest nurseries in Wisconsin proved reliable for seedbeds that were relatively uniform in seedling density. However, inconsistencies between the seedbed inventory and the number of seedlings actually lifted and shipped occurred when the seedbed density was highly variable. Such seedbed heterogeneity was attributed to the combined effects of variable seed germination, abiotic factors (wind erosion, hail damage, flooding, and mechanical damage during cultivation), and biotic factors (animal predation and localized disease and insect losses). Highly variable seedbeds required more sampling than seedbeds with uniform seedling density, but intensive sampling was too expensive. Thus, the dilemma was to find a sampling design that provided the necessary accuracy at a relatively low cost.

A simple strategy was designed and implemented to improve current inventory practices for estimating seedling numbers in a forest tree nursery operated by the Wisconsin Department of Natural Resources. We wanted to know the sample sizes needed to estimate seedling inventories in nursery seedbeds at given levels of precision and the appropriate distribution of samples across the seedbeds. We used data from an existing inventory to estimate the sample size needed. Then we tested a new sampling design on a highly variable seedbed.

Methodology

Like many bareroot nurseries, the Wilson State Forest Nursery (Boscobel, Wisconsin) produces tree seedlings in 4-ft-wide (1.22-m-wide) seedbeds several hundred feet long (figure 1a). Each seedbed contains 30 to fifty 12-ft-long (3.66-m-long) "beds"; these subdivisions are established solely for administrative purposes. Conifers are sown in 7 rows per seedbed, while hardwoods are sown in 5 rows per seedbed. Inventories are completed each summer when seedlings are 1 or 2 y old.

The current method of inventorying 2-0 conifer seedbeds (regardless of total seedbed length) consisted of counting the number of seedlings in 7 samples, 1 sample for each of the 7 rows (figure 1b). The samples were 12-ft-long segments of individual rows, distributed along an imaginary diagonal across the seedbed. Past sampling indicated that random placement of the samples usually provided better estimates, but the systematic placement of samples along a diagonal became (perhaps incorrectly) a standard method to simplify sampling. Hardwood seedlings *were* sampled in the same fashion as conifers, except that only 5 samples per seedbed were taken.

Data from a large existing inventory were used to provide preliminary estimates of the number of samples *(n)* needed for each of the 63 seedbeds at the nursery. We used the following model (Cochran 1977):

$$(1)\quad n = n_0 / [1 + (n_0 / N)]$$

where

$$(2)\quad n_0 = [(t_{\alpha/2,\, n-1} \cdot s)/(r \cdot \bar{x})]^2$$

And n= number of 12 ft long samples needed for a particular seedbed, ***N*** **=** total number of 12-ft-long row sections in the seedbed, ***t*** = tabulated value of Student t-test with x/2 ordinate (x=0.10), *n-1* degrees of freedom, s =

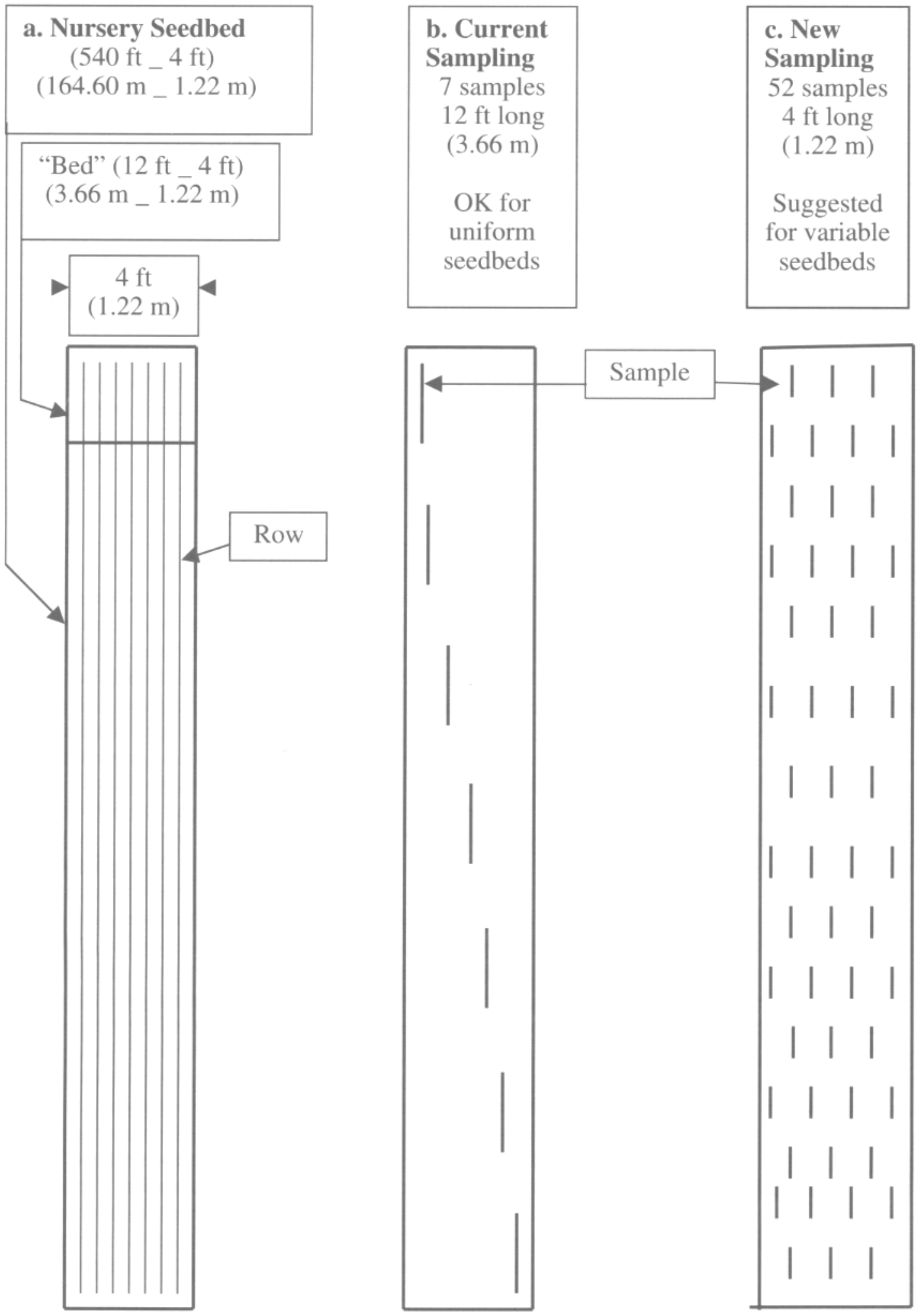

Figure 1—*Layout of (a) a conifer nursery seedbed, (b) the current sampling design used for seedling inventories, and (c) our proposed sampling design. The seedbed is not shown to scale.*

standard deviation of the sample, r = desired half width of the relative error (for example, an *r = 0.1 is* interpreted as ± 10% of the mean number of seedlings per sample), and x = mean number of seedlings per sample. An example showing the use of formulae (1) and (2) is provided in box 1.

Since large differences in seedling densities often existed among and within seedbeds, we divided the 63 seedbeds into 3 arbitrary groups based on their coefficients of variation (CV = (s / x) 100). These were identified as "best-case" (20 low-variability seedbeds, with 0 < CV <_ 20), "typical-case" (22 medium-variability seedbeds, with 20 < CV <_ 35), and "worst-case" (21 high-variability seedbeds, with 35 < CV) seedbed groups. Then we calculated the average sample size needed to provide estimates for each group of seedbeds by averaging estimated sample sizes for seedbeds within groups.

We also examined the distribution of the variability within a highly variable 2-0 red pine *(Pines resinosa* Ait.) seedbed (CV = 83) from our worst-case seedbed group. To estimate the variability among beds within a seedbed, we recorded the number of seedlings from 12ft-long samples in all 7 rows for 15 alternating beds. To estimate the variability within rows in beds, we subdivided the 12-ft-long row segments (105) into 4-ft-long segments or subsamples (315). We estimated variance components using ANOVA and VARCOMP procedures of the Statistical Analysis System (SAS 1988). The variability among 4-ft-long row segments nested within 12ft-long row segments became the error term in the analysis of variance.

Results and Discussion

The estimated sample size needed for seedbeds in the best-case group *(n= 9)* was similar to the sample size (ii = 7) currently employed with 12-ft-long (3.66-m-long) row segments along a diagonal (figure 1b). However, for the typical- *(n* = 28) and the worst-case *(n* = 75) situations, the minimum sample size needed to provide an accurate estimate of seedlings in highly variable seedbeds was well above current sampling practices. It is easy to understand why large discrepancies often existed between typical- or worst-case seedbed inventories and merchantable seedlings. Only 84 row-ft (7 samples x 12 ft) (25.6 row-m, 7 samples x 3.66 m) per seedbed had been sampled, but 336 row-ft (102.5 rowm) (28 samples x 12 ft) or *900* row-ft (274.5 row-m) (75 samples x 12 ft) would be needed for the typical- or worst-case scenarios, respectively.

Results from a highly variable (worst-case) seedbed indicated that most observed variation was distributed among beds (63% of the total variance; table 1).

Variation among rows within beds and among 4-ft-long (1.22-m-long) segments within 12-ft-long segments was essentially the same, accounting for 17% and 20% of the total variance, respectively. The relatively small variability among row segments and within row segments suggested that the size of each sample could be reduced here. This savings in time and effort would allow managers to survey a larger number of beds to account for heterogeneity in typical- and worst-case seedbeds. Thus, a sampling design using a larger number of smaller, more closely spaced samples distributed across more beds provided a better alternative for highly variable seedbeds.

Based on these results, we designed a new procedure to sample more beds using smaller sampling units. The new sampling design used fifty-two 4-ft-long row segments distributed in sets of 3 or 4 samples for each of 15 beds within a seedbed (figure 1c). Trials using this procedure produced estimates of the average number of seedlings per 4-ft-long row segment that were well within the 90% confidence interval obtained using very intensive sampling (table 2). In other words, adequate estimates of seedling numbers were achieved by measuring 208 row-ft (63.44 row-m) (52 samples x 4 ft) using the new design, rather than 336 row-ft (28 samples x 12 ft) or 900 row-ft (75 samples x 12 ft) required with the old design, for the typical- and worst-case scenarios, respectively.

We suggest the use of the proposed sampling proce-

Table 1—*Summary analysis of variance for a highly variable "worst-case" 2-0 red pine (*Pinus resinosa *Ait.) seedbed at the Wilson State Forest Nursery in Wisconsin*

Source of variation	Degrees of freedom	Mean square	% of total variance	*F* test
Beds	14	1438	63	19.8[a]
Rows (beds)	90	73	17	3.4[a]
Segments (row [bed])	210	21	20	

[a] Significant at the $P < 0.001$ level.

Table 2—*Average numbers of seedlings per 4-ft-long (1.22-m-long) row segments, and 1 standard error, based on initial intensive sampling of 63 seedbeds, in contrast to the proposed sampling design*

Estimator	Intensive sampling	Proposed sampling design	
		1st test	2nd test
Number of samples	315	52	52
Mean number of seedlings[a]	11.08 ± 0.92	11.30 ± 2.23	10.53 ± 2.41
Standard error	0.56	1.38	1.47

[a] 90% confidence interval.

dure for highly variable seedbeds. The number of samples can be adjusted proportionally to the seedbed variability for each inventoried seedbed. Sample means and standard deviations necessary to estimate the appropriate sample size (formulae 1 and 2) can be taken from existing inventories available at the nursery or from pilot samplings.

Further research is needed to test alternative procedures that might further reduce sampling time and effort without decreasing the quality of the inventory estimates (such as shorter sample segments, for example, 1-ft-long, 30.48-cm-long, row segments). Simulations to determine the optimal distribution of sampled row segments are also needed to better characterize variability among beds and within beds.

Conclusions

Highly variable nursery seedbeds require sample sizes proportional to their variability to provide accurate estimates of the number of merchantable seedlings. A large number of small samples (for example, 4-ft-long, 1.22-m-long, row segments) was preferable to a small number of large samples (for example, 12-ft-long, 3.66m-long, row segments) where considerable heterogeneity in seedling density existed. This appeared to be especially useful when most of the observed heterogeneity was among beds rather than within beds. The sample size needed to provide precise estimates of seedbed stocking is easily estimated using data from preexisting inventories or from pilot samplings.

Address correspondence to: Raymond P. Guries, University of Wisconsin-Madison, Department of Forest Ecology and Management, 1630 Linden Drive, Madison, WI 53706; **e-mail:** < rpguries@facstaff.wisc.edu >.

Acknowledgments

This research was supported by the Bureau of Forestry, Wisconsin Department of Natural Resources, and the Department of Forest Ecology and Management, University of Wisconsin-Madison. Cuauhtemoc Saenz-Romero was an Honorary Fellow at the University of Wisconsin-Madison, Department of Forest Ecology and Management, during the conduct of this work. He gratefully acknowledges the support of Centro de Investigacion en Genetica y Ambiente, Universidad Autonoma de Tlaxcala, Tlaxcala, Mexico, during his stay in the United States. Mr. Andrew Monk and the staff of the Wilson State Nursery provided excellent assistance. Dr. Rudy King and an unidentified reviewer provided a number of comments that improved the clarity and content of the paper.

References

Cochran WG. 1977. Sampling techniques. New York (NY): John Wiley & Sons. 428 p.

Steel RGD, Torrie JH. 1997. Principles and procedures of statistics: A biometrical approach. 3rd ed. New York (NY): McGraw-Hill. 666 p.

[SAS] SAS Institute Inc. 1988. SAS/STAT User's guide. Release 6.03. Cary (NC): SAS Institute. 1028 p.

Box 1—*Estimation of sample size.*

The sample size (n) needed to provide an estimate of the actual number of seedlings in a nursery seedbed depends upon the precision required and the variability within the seedbed. The formula for calculating 'n' is given by (Cochran 1977):

(1) $n = n_0 / [1 + (n_0 / N)]$

(2) $n_0 = [(t\propto_{/2,\, n-1} \cdot s) / (r \cdot \bar{x})]^2$

For this example, we used inventory data from seven 12-ft-long (3.66-m-long) row segments of a highly variable "worst-case" seedbed. The average number ($\bar{x}$) of seedlings in a 12-ft-long row was 123.4 and the standard deviation (s) was 50.5. The total number of 12-ft-long row segments (N) for that particular seedbed was 7 rows x 40 "beds" = 280. The degrees of freedom appropriate for estimating Student *t*-test was 7 - 1= 6. To be confident that 90 times in 100 the sample estimate was within a specified relative error (or error range) surrounding the population mean ($\propto = 0.10$), we consulted tabled values of Student *t*-test (for example, Steel and Torrie 1997). We chose a value (in this case, 1.943) that represented the ∝ probability (a "2-tailed" test) that the estimate would be within the relative error. A value of $r = 0.1$ was used, which provided a 20% width for the relative error (± 10% of the mean). Our experience suggested that this margin of error was acceptable because the inventory still met customer orders. To estimate the number of (12-ft-long) samples needed to be within the specified relative error containing the "true" population mean with 90% probability, we substituted into formula (2): $n_0 = [(t_{/2,\, n-1} \cdot s) / (r \cdot \bar{x})]^2 = [(1.943 \cdot 50.5) / (0.1 \cdot 123.4)]^2 = 63.23$

The formula $n = n_0 / [1 + (n_0 / N)]$ is a correction for finite population size. Substituting n_0 and N into formula (1): $n = n_0 / [1 + (n_0 / N)] = 63.23 / [1 + (63.23 / 280)] = 51.58$

Rounding the result, we concluded that a total of fifty-two 12-ft-long segments was an acceptable sample size for estimation of merchantable seedlings in that particular seedbed.