Sample size requirements for estimating biomass production in an Engelmann spruce, subalpine fir forest

by

M.A. Arthur
Department of Forestry
University of Kentucky
Lexington, KY 40546

G. Rubin
Biometrics Unit
Cornell University
Ithaca, NY 14853

BU-1155-MA

April 1992
Revised March 1994

Sample size requirements for estimating biomass production in an Engelmann spruce, subalpine fir forest

M.A. Arthur

Department of Forestry, University of Kentucky, Lexington, KY 40546

and

G. Rubin

Biometrics Unit, Cornell University, Ithaca, NY 14853

ACKNOWLEDGEMENTS: We extend our appreciation to K. Gerow, D. Reichert, and R. Paratley for valuable comments on a previous draft of the manuscript. The investigation reported in this paper (No. 93-8-187) is in connection with a project of the Kentucky Agricultural Experiment Station and is published with approval of the Director.

Sample size requirements for estimating biomass production in an Engelmann spruce, subalpine fir forest.

Abstract

Incremental growth of trees >10 cm diameter was measured in nineteen 0.04 ha forest plots (N=700 trees) in an uneven-aged subalpine forest in Colorado and growth increment was used to estimate biomass production. Taking those 700 trees as the population of interest, a simulation study was performed to determine an adequate sample size of trees for coring to provide similar estimates with less effort. We examined different stratified sampling strategies, varying the number or proportion of trees sampled from each stratum (forest plot) of the population, treating each of three size classes of tree separately. Since the total numbers of trees per size class are known constants, the basal area increment obtained from each replication of each simulation can be used to estimate the total basal area increment for each size class. Our results indicate that, at least in a forest with low species diversity, stratified random sampling using only a few trees per plot can be used to estimate basal area increment accurately.

Key words: simulation, sample size, stratified sampling, Loch Vale Watershed

INTRODUCTION

Large variation in annual growth among trees of a given size class can result in poor correlation between tree diameter and incremental growth (Oliver and Larson 1990). Hence, when researchers are interested in estimating forest biomass production they tend to measure annual increments from a large sample of the population in hopes of improving their estimates. When allometric relationships have been previously determined, forest biomass production can be estimated by coring a large fraction of the stand trees, measuring annual increments, and extrapolating to biomass production using allometric relationships for the tree species of interest (Aplet et al. 1989, Arthur and Fahey 1992, Bockheim et al. 1989, Grier et al. 1981, Nadelhoffer et al. 1985, Pastor et al. 1984). However, the inherent uncertainty in biomass estimates from allometric equations would suggest that a heavy investment in the measurement of incremental growth is not warranted. This is especially true if a high level of precision in estimated basal area increment can be achieved using a much smaller subsample of tree radial increments. In addition, the observed variation in basal area (BA) increment should be, but rarely is, used in any calculation pertaining to nutrient cycling or estimates of production; thus, the extra effort involved in calculating BA increment from a large rather than a small subsample is wasted. Without quantifying the error in estimates based on allometric equations there is little utility in expending tremendous effort in the labor-intensive task of measuring large numbers of increment cores.

Arthur and Fahey (1992) measured incremental growth of trees >10 cm diameter in nineteen 0.04 ha forest plots (N=700 trees) in an Engelmann spruce, subalpine fir forest in Colorado to obtain a measure of biomass production for use in calculating nutrient uptake. Taking those 700 trees in the 19 plots as the population of interest, we performed a simulation study, investigating several sampling strategies to determine whether a smaller number of trees could have been cored to accurately estimate total BA increment. The variance formulae for stratified sampling plans (Cochran 1977) can provide a guide in choosing among plans with a fixed number of units per stratum or with proportional allocation (i.e., the sampling fraction is the same in all strata).

However, this simulation study provides a demonstration of the behavior of several fixed sample size and proportional allocation stratified sampling plans for total BA increment, a variable which may be sensitive to the sample selected.

METHODS

Site

The study area was Loch Vale Watershed (LVWS), located in Rocky Mountain National Park, Colorado, between 3100 m and 4000 m. The forest is dominated by uneven-aged Engelmann spruce (Picea engelmannii Parry) and subalpine fir (Abies lasiocarpa (Hook.) Nutt.), which comprise 97 percent of the basal area of the forest; the remaining 3 percent is limber pine (Pinus flexilis James). Stand basal area is highly variable (range = 16.4 to 72.9 m^2 ha^{-1} for 19 plots). For a more complete description of the study site see Arthur and Fahey (1992).

Field methods

Nineteen 0.04 ha forest plots were located within the forested areas of LVWS using a stratified random sampling design to insure that all regions of the forest were included in the study. In these nineteen plots, the diameter and species of all trees greater than 10 cm dbh were recorded, and one increment core taken from each tree. There were a total of 700 increment cores, for which the last five-year radial increments were measured and extrapolations made to BA increments. The five-year BA increments were the data used in the study, the 700 trees in these 19 plots being considered as the population of interest. Hence, for the population consisting of these 700 trees, the true total BA increments of the population and the true variance of BA increment in each of the 19 strata were known (Table 1).

Simulation methods

Separate simulations were performed for each of three size classes of trees: 10-20 cm dbh, 20.1-30 cm dbh, and >30 cm dbh. We did not combine total BA increment estimates across size classes because subsampling of a forested stand for BA increment typically is done based on size class groupings (Grier and Logan 1977, Prescott et al. 1989), and because growth rate becomes more variable with increasing tree size and age, presumably due to natural senescence (Oliver and Larson 1990). The total number of trees per size class was 420, 156, and 124 for the 10-20, 20.1-30 and >30 cm dbh classes, respectively. The total number of trees per plot (N_h) ranged from 10 to 39 in the 10-20 cm class, 2 to 21 in the >20-30 cm class, and 1 to 15 in the >30 cm class.

The 19 forest plots were considered strata because each had its own environmental characteristics which were expected to influence forest growth. Hence, only stratified sampling plans were considered. Two types of stratified sampling plans were considered: those in which a simple random sample of a fixed number of trees was taken from each stratum (n_h), yielding different sampling fractions or weights for each stratum; and those with proportional allocation, yielding a common sampling fraction, f, in all strata, where f is defined as $f = (n/N) = (n_h/N_h)$. Plans with proportional allocation have the advantage of providing a self-weighting sample, and thus, simplified calculation of estimates. Four plans using a fixed number of trees per stratum (n_h=1, 2, 3 and 5) and four plans with proportional allocation (100 x n_h/N_h= 5, 10, 20 and 30 %) were tested. The convention in simulation studies is to sample with replacement; thus, in all simulations trees (sampling units) were sampled with replacement. Hence, the fixed number sampling scheme was possible even in plots with fewer than the specified number of trees to be sampled. Because some plots contained only a few trees of the larger size classes, proportional allocation plans with small sampling fractions could not be used for the 20.1-30 and >30 cm dbh size classes (i.e., with small f, these plots would not be sampled). For the 20.1-30 cm dbh size class only the 30% proportional allocation plan could be tested, and for the >30 cm dbh size class, none of the 4 proportional allocation plans could be used. For each simulation of the sampling plan for each size class of tree, 1500 replications of the sampling procedure were performed (i.e., simulation sample equals 1500).

The mean BA increment for a given size class, y, is calculated as

$$\bar{y} = \frac{\sum_{h=1}^{19} N_h \bar{y}_h}{N},$$

where N_h is the total number of trees in stratum h, $\bar{y}_h$ is the mean BA increment in stratum h and N is the total number of trees in that size class. Hence, the estimate of the population total BA increment for a given size class is:

$$\hat{Y} = N \bar{y},$$

and the variance of the estimated population total BA increment is:

$$Var(\hat{Y}) = \sum_{h=1}^{19} N_h (N_h - n_h) \frac{S_h^2}{n_h}, \qquad (1)$$

where S_h^2 is the true variance of BA increment in stratum h (Cochran, 1977). For stratified proportional allocation plans, the variance of the estimated population total BA increment is :

$$Var(\hat{Y}) = \frac{(N - n)}{n} \sum_{h=1}^{19} N_h S_h^2 \qquad (2)$$

$$= \left(\frac{1 - f}{n}\right) \sum_{h=1}^{19} (N_h / N) S_h^2, \qquad (3)$$

where n is the total number of trees sampled and f is the common sampling fraction for all plots and Nh/N is the weight of the stratum h. Estimates of (1), (2) and (3) are obtained by substituting the estimated variance in stratum h, s_h^2, for the true variance.

The estimated total BA increment is the mean of total BA increment from the 1500 replications of a simulation. The simulated standard error measures the precision of the estimate of total BA increment from the 1500 simulations; it is calculated as the square root of the variance among the 1500 estimates of total BA increment, divided by the square root of 1500. The mean standard deviation (SD) of total BA increment is the average over the 1500 replications of the square root of the estimated variance for that plan [see equations (1) and (2)].

All simulations were programed in the DATA step of SAS (SAS 1985) using 1500 replications per simulation for each size class and plan. For each replication, a simple random sample of either the fixed number of trees or the appropriate number of trees to achieve the constant sampling fraction, f, was selected from each stratum. To provide the number of the first tree to be sampled, a uniform random number between 1 and the total number of trees of that size class for that stratum was generated by multiplying a random number uniform on the interval 0 to 1 by the total number of trees for that stratum and then adding one to that product and taking the integer part. This procedure was repeated n_h times for each stratum. For each replication, the mean and variance of BA increment were calculated for each stratum and then the estimated population total BA increment and the estimated variance of the population total BA increment were calculated.

RESULTS AND DISCUSSION

Fixed sample size

The true total BA increment and its true variance are given for each stratum and size class in Table 1. Note that two plots contain only one tree in size class 3; thus, there is no estimate of variance for those two plots. Variances generally increase from size class 1 to size class 3, despite the fact that each size class spans 10 cm dbh, indicating more variable rates of growth for larger trees.

The goal of this simulation study was to determine the minimum number of trees per plot to be sampled to estimate the total BA increment reliably (with high accuracy and low variance for the estimate). For each of the three size classes, sampling one tree per stratum was sufficient for achieving an estimate of the total BA increment that was not significantly different from the true population total BA increment ($p \geq 0.99$ in all for all size classes; Tables 2-4). However, no measure of variance within and, hence, among plots can be calculated from such a plan; thus, we would recommend sampling at least 2 trees per stratum. For all the fixed sample sizes tested (n=1, 2, 3, 5 trees per plot), the estimate of BA increment was very close to the true value ($p \geq 0.98$ for all plans for each size class except n=3, size class 3). For sample sizes of 2, 3 and 5 trees per plot, the simulated standard deviation of total BA increment was close to the true standard deviation of each plan, shown by a ratio of simulated to true SD that was greater than 0.95 for all plans tested in all size classes except n=2, size class 3. This indicated that the sampling procedures were giving reasonable estimates of the variation in total BA increment for each plot. Thus, the minimum number of trees for which radial increment could have been measured was 38 trees for each size class (2 trees x 19 plots), or 114 trees instead of the 700 trees actually sampled.

Proportional sampling

We also were interested in the behavior of sampling schemes in which trees are sampled proportionally to the number per strata. For instance, because there were many more trees in the 10-20 cm size class than in the other two classes, it would be preferable to measure more trees of that size class. We considered four proportional sampling plans: 5%, 10%, 20% and 30%. For size classes 1 and 2, all proportional allocation plans tested provided an estimate of total BA increment that was not significantly different from the true population ($p \geq 0.98$ in all cases; Tables 2 and 3). Thus, proportional allocation plans provided accurate estimates for the smaller size classes of trees.

For size class 1, a 5% proportional plan (or 24 trees total) was sufficient to accurately estimate the BA increment (Table 2). This level of sampling required virtually the same effort as one tree per plot, since all but five plots had no more than one tree per plot sampled. For this size class, the precision of the estimate of total BA increment, as measured by the simulated standard error, does not change greatly with an increase in the proportion of trees sampled. However, the ratio of the simulated SD to the true SD of total BA increment was less than 0.90 for plans of less than 20 %, indicating that these plans poorly estimate the within plot variation in total BA increment. This is especially true for the 5 % proportional plan, since only 3 plots will have a variance estimate that exceeds zero. The 20 % plan, which is equivalent in total sampling effort to the 5 trees per plot fixed plan, provides a reasonable estimate of the SD of total BA increment.

For size class 2, the 30 % plan would require slightly less total sampling effort than the 3 trees per plot fixed sample scheme (Table 3). However, given that one plot has only 2 trees, one plot would be excluded from the sample, which violates the plan. Similar results are obtained with the

2 plans. For size class 3, the 30 % sampling proportion, which is similar in number of trees to 2 trees per plot, performs poorly (p=0.6621) since 5 plots are not represented. A proportional plan of greater than 50 % would be required to ensure that at least 1 tree per plot was sampled for the largest size class. Nevertheless, proportional sampling plans have the advantage of giving self-weighted estimates, and it might be worthwhile to sample proportionally even when sampling a high percentage of the trees is necessary to have all strata represented.

CONCLUSIONS

Biologists know that there is large variation in most measurable properties of a population; thus, they frequently sample very heavily, but then use allometric relationships as if they represent the truth. We make the point here that, due to the approximate nature of biomass estimates from allometric equations, a heavy investment in the sampling of trees within a stand for radial increment measurements is unnecessary, especially given the level of accuracy one can achieve with a smaller sampling effort. A researcher may actually want to know the variance of the BA increment among stands, rather than relying only on the population mean. Even so, a greatly reduced sample would still give a reasonable estimate of variance. Thus, savings of 91 % for size class 1, 76 % for size class 2 and 69 % for size class 3 can be achieved by sampling two trees per strata rather than the entire population of trees in this forest type.

This simulation study was based on increment data from a forest with very low species diversity (essentially only 2 species) and much of the uniformity within size classes can be attributed to the similarity in growth rates within a size class for the two species. Thus, the applicability of this study is constrained to forests with similarly low species diversity.

LITERATURE CITED

Aplet, G.H., Smith, F.W. and Laven, R.D. 1989. Stemwood biomass and production during spruce-fir stand development. J. Ecology 77:70-77.

Arthur, M.A. and Fahey, T.J. 1992. Biomass and nutrients in an Engelmann spruce, subalpine fir forest in north-central Colorado: pools, annual production and internal cycling. Can. J. For. Res. 23:315-325.

Bockheim, J.G., Leide, J.E., and Frelich, L.E. 1989. Red pine growth and chemical composition of foliage and forest floors across a precipitation-chemistry gradient in Wisconsin. Can. J. For. Res. 19:1543-1549.

Cochran, W.G. 1977. *Sampling Techniques, Third edition.* John Wiley and Sons, New York, 428 pp.

Grier, C.C. and Logan, R.S. 1977. Old-growth *Pseudotsuga menziesii* communities of a western Oregon watershed: Biomass distribution and production budgets. Ecological Monographs 47:373-400.

Grier, C.C., Vogt, K.A., Keyes, M.R., and Edmonds, R.L. 1981. Biomass distribution and above- and below-ground production in young and mature *Abies amabilis* zone ecosystems of the Washington Cascades. Can. J. For. Res. 11:155-167.

Nadelhoffer, K.J., Aber, J.D. and Melillo, J.M. 1985. Fine roots, net primary production, and soil nitrogen availability: A new hypothesis. Ecology 66:1377-1390.

Oliver, C.D. and Larson, B.C. 1990. *Forest Stand Dynamics.* McGraw Hill, Inc., NY.

Pastor, J., Aber, J.D. and McClaugherty, C.A. 1984. Aboveground production and N and P cycling along a nitrogen mineralization gradient on Blackhawk Island, Wisconsin. Ecology 65:256-268.

Prescott, C.E., Corbin, J.P. and Parkinson, D. 1989. Biomass, productivity, and nutrient-use efficiency of aboveground vegetation in four Rocky Mountain coniferous forests. Can. J. For. Res. 19:309-317.

Sas Institute Inc. 1985. *Sas User's Guide: Version 5 edition.* Cary, NC.

Table 1: True total basal area (BA) increment (cm^2) and true variance (Sh^2) by strata (plot) and size class (10-20 cm dbh, 20.1-30 cm dbh, and >30 cm dbh) for the population of 700 trees. The number of trees per stratum (N_h) are given for each size class.

Strata (h)	Tree Size Class								
	10-20 cm			20.1-30 cm			> 30 cm		
	N_h	Total BA increment	Variance (S_h^2)	N_h	Total BA increment	Variance (S_h^2)	N_h	Total BA increment	True (S_h^2)
1	16	160.7	68.9	9	85.6	33.5	8	151.3	244.5
2	26	293.3	36.5	12	173.1	173.6	6	97.8	48.9
3	22	208.8	42.1	6	111.9	65.1	1	27.0	*
4	10	87.8	28.6	12	224.7	83.5	6	133.9	136.8
5	17	216.4	94.9	2	52.6	89.8	4	65.4	193.5
6	24	340.2	70.7	6	99.2	119.4	8	228.8	234.9
7	36	370.0	92.4	11	167.5	73.7	2	83.8	32.0
8	17	260.6	65.6	6	154.4	112.9	6	245.0	403.1
9	13	168.0	86.1	6	86.7	73.0	8	229.1	242.9
10	38	353.6	60.0	3	59.2	135.3	8	112.8	92.0
11	38	426.5	49.9	13	355.4	162.3	1	97.2	*
12	18	164.1	39.3	6	98.2	192.4	15	241.4	146.0
13	14	230.2	67.4	6	89.5	84.2	10	162.1	155.0
14	39	512.1	65.2	21	631.7	186.2	2	139.0	15.7
15	13	181.4	102.9	7	239.9	218.2	4	300.9	417.2
16	33	323.5	40.0	11	174.7	54.2	15	231.8	221.0
17	11	121.7	53.7	10	166.6	85.5	13	419.4	846.4
18	11	174.4	66.2	5	119.6	150.0	5	167.5	214.5
19	24	245.5	60.9	4	81.9	232.2	2	49.5	13.0
	N=420			N=156			N=124		

* Cannot be calculated because N=1 tree per plot.

Table 2: Results of simulation study for trees 10-20 cm dbh (size class 1). True population total basal area (BA) increment for this size class was 4838.8 cm^2. The number of replications for each simulation was 1500. The true standard deviation (SD) of total BA for each plan was calculated from the data for the whole population and, hence, was a known constant. The mean SD of total BA for each plan is the average of the estimated standard deviation of total BA from each of the 1500 replications for that plan.

Sampling Plan	Total # of trees sampled	Estimated total BA increment (sim. SE)	True SD of plan	Mean SD of plan based on 1500 reps	Ratio of mean SD to true SD	Z	p-value+ (2-tailed)
# trees/plot							
1	19	4842.6 (20.9)	809.7	*	*	0.0047	0.9963
2	38	4836.3 (15.0)	561.1	554.5	0.9882	0.0045	0.9964
3	57	4835.1 (12.0)	448.6	441.0	0.9831	0.0083	0.9934
5	95	4845.7 (9.1)	332.2	327.4	0.9856	0.0207	0.9835
Proportion							
5%	24	4834.5 (17.7)	702.2	456.0	0.6494	0.0062	0.9951
10%	42	4841.2 (13.1)	483.3	425.2	0.8798	0.0049	0.9961
20%	84	4843.9 (13.1)	322.2	317.4	0.9851	0.0158	0.9874
30%	126	4843.9 (7.1)	246.1	233.6	0.9492	0.0193	0.9846

* Cannot be calculated because n=1 tree per plot.
\+ p-value = probability of observing a value of the Z-statistic as large or larger than that actually observed.

Table 3: Results of simulation study for trees 20.1-30 cm dbh (size class 2). True population total basal area (BA) increment for this size class was 3172.4 cm^2. The number of replications for each simulation was 1500. The true standard deviation (SD) of total BA for each plan was calculated from the data for the whole population and, hence, was a known constant. The mean SD of total BA for each plan is the average of the estimated standard deviation of total BA from each of the 1500 replications for that plan.

Sampling Plan	Total # of trees sampled	Estimated total BA increment (sim. SE)	True SD of plan	Mean SD of plan based on 1500 reps.	Ratio of mean SD to true SD	Z	p-value+ (2-tailed)
# trees/plot							
1	19	3169.8 (11.6)	443.8	*	*	0.0059	0.9953
2	38	3173.0 (8.2)	298.0	280.0	0.9396	0.0020	0.9984
3	57	3176.9 (6.7)	229.7	220.0	0.9578	0.0196	0.9844
5	95	3171.9 (5.1)	154.7	149.9	0.9690	0.0032	0.9975
Proportion							
30%	47	3174.8 (6.0)	211.8	193.0	0.9112	0.0113	0.9910

* Cannot be calculated because n=1 tree per plot.

\+ p-value = probability of observing a value of the Z-statistic as large or larger than that actually observed.

Table 4: Results of simulation study for trees >30 cm dbh (size class 3). True population total basal area (BA) increment for this size class was 3183.7 cm^2. The number of replications for each simulation was 1500. The true standard deviation (SD) of total BA for each plan was calculated from the data for the whole population and, hence, was a known constant. The mean SD of total BA for each plan is the average of the estimated standard deviation of total BA from each of the 1500 replications for that plan.

Sampling Plan	Total # of trees sampled	Estimated total BA increment (sim. SE)	True SD of plan	Mean SD of plan based on 1500 reps.	Ratio of mean SD to true SD	Z	p-value+ (2-tailed)
Trees/plot							
1	19	3177.0 (14.1)	545.9	*	*	0.0123	0.9902
2	38	3192.7 (10.0)	364.9	347.0	0.9509	0.0247	0.9803
3	57	3170.7 (8.2)	297.6	269.2	0.9628	0.0465	0.9629
5	95	3183.2 (6.3)	185.0	177.3	0.9584	0.0027	0.9979

* Cannot be calculated because n=1 tree per plot.

+ p-value = probability of observing a value of the Z-statistic as large or larger than that actually observed.