

Estimation of seed orchard efficiencies by means of multistage variable probability sampling

C. BARTRAM

British Columbia Ministry of Forests, Silviculture Branch, 1450 Government Street, Victoria, B.C.,
Canada V8W 3E5

AND

G. MILLER

Canadian Forestry Service, Pacific Forestry Centre, 506 West Burnside Road, Victoria, B.C., Canada V8Z 1M5

Received June 30, 1987

Accepted July 11, 1988

BARTRAM, C., and MILLER, G. 1988. Estimation of seed orchard efficiencies by means of multistage variable probability sampling. *Can. J. For. Res.* **18**: 1397-1404.

This paper describes two frameworks for developing effective multistage variable probability sampling methods to estimate seed orchard efficiencies. In both cases a standard multistage approach is implemented initially in multiple orchards and years. The effectiveness of this approach is then evaluated against several alternative methods, using the initial efficiency data collected, and subsequent survey procedures are prescribed on the basis of this analysis. An example illustrating the application of this methodology in coastal Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) seed orchards in British Columbia is presented.

BARTRAM, C., et MILLER, G. 1988. Estimation of seed orchard efficiencies by means of multistage variable probability sampling. *Can. J. For. Res.* **18** : 1397-1404.

Les auteurs décrivent deux modèles pour l'élaboration de méthodes efficaces d'échantillonnage à divers degrés et avec probabilités variables pour l'estimation du rendement des vergers à graines. Dans les deux cas, une démarche normalisée d'échantillonnage à divers degrés est initialement appliquée dans plusieurs vergers et pour de nombreuses années. L'efficacité de cette démarche est ensuite évaluée par rapport à nombre d'autres méthodes à partir des données sur le rendement qui ont été obtenues initialement, et les méthodes subséquentes d'échantillonnage sont déterminées à partir de cette analyse. Un exemple d'application de cette démarche dans des vergers à graines de Sapin de Douglas côtiers (*Pseudotsuga menziesii* (Mirb.) Franco) en Colombie-Britannique est présenté.

Introduction

In conifers, the potential number of seeds for a crop is generally determined by the number of female strobili initiated, cone survival, and the number of fertile scales per cone. The number of filled seeds per cone is determined by the number of fertile scales, the level of pollination, the number of seeds lost to insects, and the rates of seed abortion and malformation. Cone mortality rates are variable and often high (e.g., Silen and Keane 1969; White *et al.* 1977; Zasada 1971). Likewise, conifer seed yields vary considerably (Anonymous 1974; Dobbs *et al.* 1976).

Because levels of seed production in southern pine (*Pinus* spp.) seed orchards were often disappointing, a procedure was developed for evaluating production efficiencies (Bramlett *et al.* 1977; Bramlett and Godbee 1982). This procedure allowed for the evaluation for seed production as affected by orchard management practices and the identification and quantification of causes of seed loss that occurred at various times.

Effective survey procedures have likewise been developed for estimating efficiencies in British Columbia's orchard program. Efforts have focused primarily on the estimation of cone efficiency (CE) and seed efficiency (SE), defined as proportions by Bramlett and Godbee (1982) as follows:

$$CE = \frac{\text{number of cones harvested}}{\text{number of conelets initiated}}$$

$$SE = \frac{\text{number of filled seeds}}{\text{number of fertile sites}}$$

and mean filled seed per cone (FS). Initial efficiency estimates were obtained via one of two multistage approaches, depending on which of these variables was of interest. In

each case, alternative multistage methods were than evaluated and effective procedures for subsequent surveys were prescribed.

This methodology is presented because it may provide a suitable framework for orchardists elsewhere who are faced with the task of estimating these efficiencies. An example showing the application of these procedures for estimating these variables in Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) seed orchards in coastal British Columbia is also presented.

The initial and alternative multistage methods described are proposed for several reasons. Sample selection procedures are straightforward. Furthermore, equations for estimating means and variances are not complex. As a result, the expertise required to implement any of these methods is not beyond orchardists with only limited exposure to survey sampling. Sample selection time is also short, because efforts are focused on units drawn at the first stage.

It is assumed in this paper that single, overall mean estimates are required annually for each orchard.

Initial estimation

Case I: CE, and SE or FS, of interest

The initial approach

For each species of interest it is recommended that at least two orchards and 2 years be evaluated initially, to enable an appraisal of site and annual variation to be made. Independent sampling is undertaken, as follows, in each orchard and year.

Just after the time of flowering, each tree in the orchard is observed and an ocular estimate is made of the number

TABLE 1. Example showing selection range compilation for individual trees

Tree	Conelet estimate (M_i)	Cumulative sum (ΣM_i)	Selection range
1	100	100	1 - 100
2	0	100	—
3	200	300	101 - 300
4	100	400	301 - 400
•	•	•	•
•	•	•	•
•	•	•	•
N-1	150	49 250	- 49 250
N	500	49 750 (= M_0)	49 251 - 49 750

of female conelets present (M_i). At least 40 trees are then selected, with probability proportional to estimated size of crop, i.e., variable probability selection. One way of achieving such a selection is to first form selection ranges for individual trees on the basis of cumulative sums of the M_i values, as shown in Table 1. Numbers are then drawn using a table of random numbers between 1 and the observed crop size over all (N) trees in the orchard (M_0). A tree is selected whenever a number falls within its selection range. The random numbers are always selected with replacement, so that the tree selection probabilities remain at M_i/M_0 with each draw.

In large orchards, especially, the effort required to estimate M_i and compile a selection range for each tree may be prohibitive. A quicker procedure is to assign a conelet rating (e.g., light, heavy, etc.) to each tree and on the basis of the average conelet estimate per tree within each rating (\bar{M}_i) and the corresponding number of trees so rated (N_i), derive a selection range for each group of trees with the same rating (Table 2). A tree is selected by first selecting a group by means of variable probability selection and then selecting a specific tree within the group at random. Sampling is again done with replacement in both instances. Less time is required because trees can be evaluated more quickly and selection ranges need only be compiled for each group, rather than for each tree.

These two methods would be equivalent in the event that all trees, if individually assigned an M_i value, were consistently given the group mean values, i.e., M_i is always equal to \bar{M}_i . In this situation the tree selection probabilities remain the same:

$$\frac{M_i}{M_0} = \frac{\bar{M}_i N_i}{M_0} \times \frac{1}{N_i}$$

Each selected tree is revisited and subsampled as follows. Ocular estimates of M_i are made for each productive whorl and at least one whorl is then sampled using variable probability selection. If a well-defined whorl structure is not present, crown levels may alternatively be used. One or more branches within each selected whorl are then sampled, again using variable probability selection, following the assignment of individual branch M_i values within the whorl. Selection of multiple whorls per tree and multiple branches per whorl is desirable, as this will enable an appraisal of different subsampling allocations to be made. Such multiple-unit subsampling is again done with replacement.

This within-orchard variable probability selection of whorls and branches can be made more practical by employing a hand-held computer which compiles the selection ranges as the M_i value of each whorl or branch is entered and then performs the variable probability selection accordingly. Software written for this purpose, for use with a Husky computer, is available from the authors on request. Sampling with equal probabilities via, say, the drawing of numbered tags from a small bag could also be considered. Such a practice will, however, introduce bias unless the CE, SE, and FS values of the individual whorls and branches are not correlated with the corresponding M_i values (Sukhatme and Sukhatme 1970, p. 342).

When each branch is selected, it is flagged to ensure easy relocation and the total number of cones on the branch is tallied. If large numbers of conelets are present a tag could be placed along each branch and only those conelets between it and the branch tip counted. This practice will not introduce bias, providing the CE, SE, and FS values of the excluded conelets do not differ from the values for those sampled.

Recounts are subsequently made just before cone harvest to enable cone efficiency to be estimated. If the timing of cone abortion losses is of interest, additional recounts will be needed throughout the season. During the final recount, one or more cones are sampled randomly from the selected branches, again with replacement if allocation of multiple cones per branch is employed.

Each sampled cone is first sliced and the number of filled seeds severed on one cut surface is tallied. A slice orientation that yields the highest correlation between the severed and total filled seed per cone is preferable. For many species, such as Douglas-fir, a longitudinal slice along the cone axis (Winjum and Johnson 1960) is best. However, for some species, such as lodgepole pine (*Pinus contorta* var. *latifolia* Engelm.) and western red cedar (*Thuja plicata* Donn), a cross-sectional slice has proved superior (G. Miller, unpublished data). Each sliced cone is then fully dissected scale by scale and the total number of filled seeds and fertile scales is determined. Judgment must be used in deciding whether or not a scale is fertile. This decision is based on the occurrence of a full-sized seed or the location of the scale within the cone and its morphology. For most conifers the number of fertile sites within a cone equals the number of fertile scales multiplied by 2.

Throughout this sample selection and evaluation process it is important that the cost of each sample selection and evaluation step be determined. This information will be needed to evaluate the effectiveness of this approach against several alternatives.

In summary, as illustrated in Fig. 1, this case I initial approach consists of three- and four-stage variable probability sampling methods for CE and SE or FS estimation, respectively. Trees are sampled first (stage 1). Whorls, or crown levels (stage 2), followed by branches (stage 3) are then selected for CE estimation. Lastly, cones (stage 4) are sampled for SE or FS estimation. Sampling of multiple units at each stage is always done with replacement, and each time a unit is drawn repetitively a new and independent subsampling is done at all subsequent stages. Under this scheme the general multistage mean estimator (\bar{Y}_{ms}) for each orchard and year evaluated is

TABLE 2. Example showing selection range compilation for groups of trees

Tree rating	Average conelet estimate per tree (\bar{M}_i)	No. of trees (N_i)	Conelet estimate per group ($\bar{M}_i N_i$)	Cumulative sum ($\sum \bar{M}_i N_i$)	Selection range
Nil	0	25	0	0	—
Light	100	50	5 000	5 000	1 - 5000
Medium	300	100	30 000	35 000	5 001 - 35 000
Heavy	500	50	25 000	60 000 (= M_0)	35 001 - 60 000

TABLE 3. Initial and alternative multistage methods for each case and variable (with last-stage unit)

Variable	Method	
	Initial	Alternative
Case I		
CE	Three-stage (branch)	Four-stage (conelet)
SE	Four-stage (cone)	Five-stage (fertile site)
FS	Four-stage (cone)	Four-stage, with two phases at last stage (cone)
Case II		
SE	Two-stage (cone)	Three-stage (fertile scale)
FS	Two-stage (cone)	Two-stage with two phases at last stage (cone)

NOTE: CE, cone efficiency; SE, seed efficiency; FS, filled seed per cone.

$$[1] \quad \bar{Y}_{ms} = \frac{1}{n} \sum_i^n \bar{Y}_i$$

where n is the number of trees selected and \bar{Y}_i is the mean CE value of all branches, or mean SE or FS value of all cones, sampled from the i th tree. When the same number of branches or cones is selected per tree, this formula is simplified to that given by Sukhatme and Sukhatme (1970, pp. 341-342), i.e., a simple arithmetic mean of the individual branch or cone values.

This estimator is unbiased if the M_i values assigned at each stage are perfectly correlated with the true M_i values. Within practical limitations, efforts should therefore be made to ensure that a reasonably good relationship is achieved. The cone selection procedure can be criticized on this ground because the M_i values at the first three stages were derived from numbers of *conelets*, not cones, observed. For example, a tree that experiences high conelet abortion would receive an M_i that exceeds its true cone-based value. However, this procedure is proposed for practical reasons because it allows the estimation of SE or FS without the necessity of repeating the sample unit selection process once the cones are mature. The random sampling of cones at the last stage for SE estimation introduces a further approximation, because cone selection probabilities should, in theory, be dependent on the number of fertile sites per cone. This is of little consequence whenever the number of fertile sites per cone is about the same.

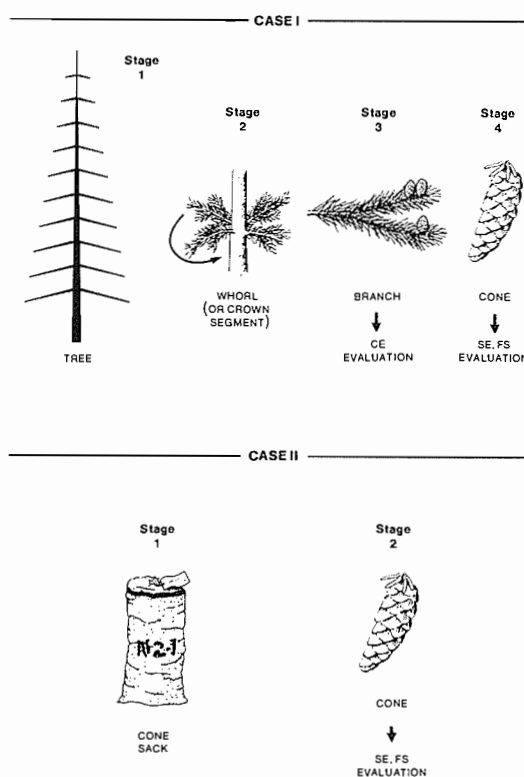


FIG. 1. The stages of sampling for each case.

Much of this initial approach is consistent with the comprehensive guidelines given by Bramlett and Godbee (1982). Some differences exist, however, which may result in better estimates. Use of variable rather than equal probability selection, for example, will reduce bias in the event that eq. 1 is employed and a correlation exists between M_i at any stage and CE, SE, or FS values. The selection of whorl and then branch within a tree, rather than branch directly, is also easier because fewer branches need be considered when the sample is drawn. This will also help to ensure that the selected branches are well distributed throughout the productive portion of the tree crowns. The sampling of new trees each year is more generally applicable, as many species do not bear a crop consistently each year.

Alternative approaches

The three- and four-stage methods implemented initially are not necessarily the most effective. Alternative subsampling allocations may yield estimates of specified precision at less cost. Other multistage variable probability sampling methods such as those listed for case I in Table 3 are also feasible and may likewise prove superior. For CE,

the alternative four-stage method would entail further selection of one or more conelets from each sampled branch; for SE, the five-stage method would similarly require the sampling of fertile sites from each sampled cone. The addition of these stages may prove worthwhile if any cost reductions resulting from increased precision and reduced branch or cone evaluation time are not offset by the added cost of processing a sample that is more evenly distributed throughout the orchard. CE and SE estimates for both these alternative methods would again be obtained using eq. 1.

In the two-phase system for FS, the four-stage method would be used to sample cones, each of which (phase 1) would be sliced and the number of filled seed severed per cone (X_i) recorded. A randomly selected subsample, without replacement (phase 2), would also be fully dissected and the total filled seeds per cone (Y_i) determined. A linear, least-squares regression relating Y_i to X_i is calculated from the phase 2 sample and the estimator (\bar{Y}_{ms}) is

$$[2] \quad \bar{Y}_{ms} = \bar{Y} + B(\bar{X}' - \bar{X})$$

where \bar{Y} is the dissection count mean, \bar{X}' and \bar{X} are mean slice counts of the phase 1 and 2 samples, and B is the regression coefficient (Cochran 1977, p. 339). This approach will become more attractive as both the correlation between X_i and Y_i and the ratio of the cost of dissecting versus slicing a cone increase. This estimator will, however, be biased if a linear relationship between X_i and Y_i does not exist.

Stratification could be considered for any of the case I methods in Table 3. Before selection, trees would first be divided into one of L nonoverlapping subpopulations which encompass, in total, the entire orchard. A multistage sample is then drawn independently from each stratum and the corresponding estimator (\bar{Y}_{ms}) is

$$[3] \quad (\bar{Y}_{ms}) = \sum_{h=1}^L W_h \bar{Y}_{ms(h)}$$

where W_h and $\bar{Y}_{ms(h)}$ are the proportion of the crop and mean estimate in the h th stratum, respectively. Separation of trees by phenology of female bud flush, crop size, or position within orchard, for example, could prove beneficial. In general, benefits from stratification are most probable when the within-strata variation is low. This evaluation should therefore focus, at least initially, on instances where differences among strata means are most pronounced. If worthwhile cost savings are found, then other less promising instances should also be examined.

Approach evaluation

The effectiveness of the three- and four-stage methods implemented initially can be generally evaluated against these alternative approaches for each orchard and year as follows. Design effect (D) values are first calculated:

$$[4] \quad D = \frac{v(\bar{Y}_{ms})}{v(\bar{Y}_{srs})}$$

where $v(\bar{Y}_{ms})$ is the estimated variance of the multistage CE, SE, or FS estimate and $v(\bar{Y}_{srs})$ is the corresponding variance, using simple random sampling with an equivalent number of sampling elements, i.e., conelets, fertile sites, or cones, respectively. D values provide an estimate of the number of times more elements required to yield an equally precise CE, SE, or FS estimate, using multistage sampling rather than simple random sampling.

The number of first-stage units (trees, in this case) required for each multistage method, n_1 , is next calculated as follows:

$$[5] \quad n_1 = \frac{Dn_{srs}}{\bar{m}}$$

where n_{srs} is the sample size required for simple random sampling and \bar{m} is the mean number of elements subsampled per tree. Equations to derive n_{srs} for a specified margin of error that is either absolute or relative to the mean are readily available (e.g., Cochran 1977, pp. 75-78). Once n_1 has been calculated, the implementation cost (i.e., $n_1 \times$ cost per tree) of each method can be derived and used as a basis for evaluation.

Appropriate equations for deriving the variance estimates in eq. 4 are as follows. For the three- and four-stage methods implemented (Sukhatme and Sukhatme 1970, p. 342):

$$[6] \quad v(\bar{Y}_{ms}) = \frac{\sum_i (\bar{Y}_i - \bar{Y}_{ms})^2}{n(n-1)}$$

where \bar{Y}_i , \bar{Y}_{ms} , and n are defined as in eq. 1. This is simply the estimated among-tree variance divided by n . For the proportions CE and SE (Cochran 1977, p. 52):

$$[7] \quad v(\bar{Y}_{srs}) = \frac{\bar{Y}_{ms}(1 - \bar{Y}_{ms})}{n-1}$$

where \bar{Y}_{ms} is defined as in eq. 1 and n is now the corresponding number of conelets or fertile sites sampled, respectively, to derive $v(\bar{Y}_{ms})$. For FS, $v(\bar{Y}_{srs})$ can be calculated using a subset of the data consisting of a single cone per tree for each time the tree was selected. This allocation is equivalent to simple random sampling, and eqs. 1 and 6 therefore simplify to the equations for this method when sampling is done with replacement (Cochran 1977, pp. 22, 26), i.e.,

$$\bar{Y}_{ms} = \Sigma Y_i / n$$

$$v(\bar{Y}_{ms}) = \frac{\Sigma (Y_i - \Sigma Y_i / n)^2}{n(n-1)}$$

For the alternative four-stage method, to estimate CE, $v(\bar{Y}_{ms})$ is again calculated using eq. 6; however, only a subset of the data, consisting of one or more conelets per branch, is employed. The corresponding value $v(\bar{Y}_{srs})$ is derived, again using eq. 7, but with n adjusted to reflect the reduced number of conelets used in $v(\bar{Y}_{ms})$. Variance estimates for the five-stage method for SE are calculated similarly, using data that comprise one or more fertile sites per cone only.

For the alternative two-phase system to estimate FS:

$$[8] \quad v(\bar{Y}_{ms}) = \frac{S_{y \cdot x}^2}{n} + \frac{S_y^2 - S_{y \cdot x}^2}{n'} - \frac{S_y^2}{n'N}$$

where $S_{y \cdot x}^2$ is the residual variance from the Y_i on X_i regression, S_y^2 is the variance of Y_i , n' and n are the numbers of cones in the phase 1 and 2 samples, and N is the total number of cones in the orchard. This equation is valid when cones are sampled randomly, and data for a single cone per tree should therefore be employed for deriving $S_{y \cdot x}^2$. Equation 8 is equivalent to that of Cochran (1977, p. 343), except that the last term on the righthand side is modified because a cone in the phase 1 sample can be selected repetitively. Cochran gives a further adjustment if

1/n is not negligible with respect to 1. The value $v(\bar{Y}_{srs})$ is as calculated previously for FS estimation, using the four-stage method implemented.

The optimal proportion of cones dissected (P_d) in eq. 8 (Cochran 1977, p. 341) is

$$[9] \quad P_d = \sqrt{\frac{c' (1 - \rho^2)}{c \rho^2}}$$

where c' and c are the costs of slicing and dissecting a cone, respectively, and ρ is the correlation coefficient between Y_i and X_i .

When stratification is employed, the general variance estimator is

$$[10] \quad v(\bar{Y}_{ms}) = \sum_h^L W_h^2 v(\bar{Y}_{ms(h)})$$

where L and W_h are defined as in eq. 3 and $v(\bar{Y}_{ms(h)})$ is the estimated variance of the efficiency estimate for stratum h . If the costs of sampling and evaluating a branch or cone are comparable across strata, the following optimal allocation, analogous to Neyman allocation for stratified random sampling, should be used in eq. 10:

$$[11] \quad n_h \propto W_h \sqrt{S_h^2}$$

where n_h is the sample size in stratum h and S_h^2 is the estimated variance among trees in stratum h .

Alternative subsampling allocations can be appraised most simply by means of the same calculations of D , n_1 , and implementation costs. Variance estimates for D are calculated by means of the same equations, but using subsets of the data that consist of the specific allocations of interest; n for $v(\bar{Y}_{srs})$ in the calculation of D is reduced accordingly.

Case II: only SE or FS of interest

The initial approach

If estimates of SE or FS only are required, the need to select and subsample trees for CE estimation no longer exists. In this case, cones can be more conveniently sampled following harvest from the cone collection sacks. As before, multiple orchards and years should be initially evaluated as follows.

An ocular estimate of the cone volume within each collection sack, M_i , is made first, and at least 40 sacks are sampled using variable probability selection (i.e., sack selection probability = M_i/M_0 , where M_0 is $\sum M_i$ across all sacks) and replacement. Multiple cones are then sampled at random from each selected sack. For practical reasons, this subsampling can be achieved by simply reaching into the sack and drawing representative cones. Each selected cone is sliced and dissected as described previously. The cost of each step needs to be determined again to allow an evaluation of alternative approaches to be made.

The case II procedure, as illustrated in Fig. 1, comprises two stages consisting of sacks (stage 1) and then cones within sacks (stage 2). Under this scheme the mean estimator in eq. 1 applies for both SE and FS, where n is now the number of sacks sampled and \bar{Y}_i is the mean SE or FS of all cones sampled from the i th sack.

Alternative approaches

This two-stage approach may again not be the most effective. Alternative multistage methods exist, and are listed in Table 3. For SE, fertile sites could be further sampled from

each cone, as in case I; for FS, the two-phase approach could be adopted. The mean estimators in equations 1 and 2, respectively, would apply. Stratification and selection of sacks by time of harvest may also prove worthwhile. If the cones open in the sacks soon after harvest, such a stratification may be required solely for biological reasons.

Each of these methods can be evaluated by calculating D , n_1 (i.e., number of sacks), and cost, as described for case I. The $v(\bar{Y}_{ms})$ values required to calculate D for the two-stage method implemented and alternative three-stage method for SE are derived by means of eq. 6, and for the two-phase approach by means of eq. 8, with S_y^2 calculated using data for single cones per sack (for each time a sack was selected). Equation 10 again applies if sacks are stratified by time of harvest. In each instance the corresponding value of $v(\bar{Y}_{srs})$ in D is calculated as in case I, again using a number of fertile sites or cones that is always equivalent to that used in $v(\bar{Y}_{ms})$. Alternative sack subsampling allocations can also be appraised using this technique.

Sampling of sacks *without* replacement may yield implementation cost improvements, especially as n approaches the total number of sacks filled. This alternative approach is, however, deemed to be less practical because sampling procedures and $v(\bar{Y}_{ms})$ equations are more complex. Any reduction in the among-sack contribution to $v(\bar{Y}_{ms})$ will also be diluted by the within-sack contribution, which remains unchanged. If each sack contains the same cone volume, procedures and equations for $v(\bar{Y}_{ms})$, as given by Cochran (1977, pp. 278-279), are more straightforward.

Subsequent estimation

Results from the evaluation of the initial and alternative approaches will help to guide subsequent survey procedures. Obviously, the methods and allocations that yielded low implementation costs for the specified margin of error should be adopted. However, cost savings should be weighed against any added implementation difficulties. Consistency of performance across orchards and years should also be considered. One would not want to prescribe an approach that had done well on occasion only.

Generally, if the selection of multiple units at the *last* stage is prescribed, these units may be sampled *without* replacement without introducing bias. Any resultant reductions in $v(\bar{Y}_{ms})$, however, will likely be slight, because sampling fractions will be small and (or) the contribution of all previous stages to $v(\bar{Y}_{ms})$ will remain the same. Last-stage units, selected without replacement, will remain eligible for reselection if units at any previous stage are drawn repetitively.

Once an approach has been prescribed, sample size requirements can be evaluated by calculating the corresponding n_1 values, by means of eq. 5, for each orchard and year, using an appropriate margin of error. The margin of error adopted will depend on the survey objectives. If, for example, the categorization of orchard performance (e.g., good, fair, etc.) is of interest, an absolute error, independent of the mean, should be appropriate. Conversely, if low efficiency values need to be estimated with greater precision, an error relative to the mean would be in order. If the resultant n_1 values are consistent, it is possible to prescribe a general sample size that is likely to be adequate and will not

TABLE 4. Seed orchards evaluated in the illustrative example

Orchard	Location	Establishment date	Composition	No. of trees	Orchard area (ha)
Dewdney	Saanich	1975	OP families + clones	2801	7.2
Koksilah	Duncan	1970-1972	OP families	541	3.2
Quinsam	Campbell River	1963	Clones + OP, CP families	1362	6.6
Snowdon	Campbell River	1971	Clones + OP, CP families	1074	4.8

NOTE: OP, open pollinated; CP, control pollinated.

TABLE 5. Estimates of CE, SE, and FS (\pm standard error) for each orchard and year

Orchard	Year	CE	SE	FS
Dewdney	1983	0.62 \pm 0.027	0.52 \pm 0.017	38.6 \pm 1.41
	1984	0.10 \pm 0.026	0.22 \pm 0.009	19.2 \pm 0.82
	1985	0.47 \pm 0.044	0.31 \pm 0.019	26.1 \pm 1.73
Koksilah	1983	0.70 \pm 0.026	0.37 \pm 0.019	28.3 \pm 1.29
	1984	0.34 \pm 0.050	0.06 \pm 0.099	4.6 \pm 0.48
	1985	0.75 \pm 0.040	0.32 \pm 0.016	22.0 \pm 1.18
Quinsam	1983	0.87 \pm 0.019	0.56 \pm 0.017	41.6 \pm 1.47
	1984	0.22 \pm 0.042	0.16 \pm 0.009	11.6 \pm 0.74
	1985	0.52 \pm 0.049	0.38 \pm 0.010	27.1 \pm 1.30
Snowdon	1983	0.77 \pm 0.025	0.41 \pm 0.018	28.6 \pm 1.40
	1984	0.16 \pm 0.039	0.10 \pm 0.007	6.4 \pm 0.44
	1985	0.20 \pm 0.038	0.34 \pm 0.019	25.7 \pm 1.66

NOTE: Under optimal controlled pollination conditions, FS values as high as 45-50 (ca. 0.65 SE) have been achieved for coastal Douglas-fir (C. Heaman, personal communication). For explanation of abbreviations see Table 3.

result in much wasted effort. Conversely, if n_1 values are variable, pilot studies should be initiated before completing the full survey to gauge the requirement in each specific instance.

Sample sizes calculated using eq. 5 are appropriate when single CE, SE, or FS estimates only are needed for an overall orchard each year. Should estimates of equal precision be needed for subdivisions within the orchard (e.g., for individual clones or blocks), the sampling effort required will increase. In general, the total number of first-stage units required for all k subdivisions of interest (n'_i) will be

$$[12] \quad n'_i = \frac{\sum_{i=1}^k (D_i n_{i(srs)})}{\bar{m}_i}$$

If the D_i , $n_{i(srs)}$, and \bar{m}_i values for each subpopulation are not very different from those of the entire orchard, n'_i will be about k times that required for the single, overall estimate.

Illustrative example

The following example demonstrates the application of case I methodology to coastal Douglas-fir seed orchards in British Columbia.

Specific methods

Four orchards, located on eastern Vancouver Island (Table 4), were evaluated initially over a 3-year period. Within each orchard, 150, 90, and 50 trees were selected in

1983, 1984, and 1985, respectively. A single whorl, then a branch, was selected from each tree each year for CE estimation. One cone in 1983 and up to three cones per branch in 1984 and 1985 were then selected for SE and FS estimation.

Mean and variance estimates were calculated using eqs. 1 and 6 for each orchard and year. The potential of all case I alternative methods in Table 3 was evaluated. Specifically, allocation of a single cone per branch and a fertile scale per cone was employed for the alternative four- and five-stage methods for CE and SE, respectively. This is equivalent to simple random sampling, and therefore D values always equalled 1.0.

Stratification by phenological group was considered in 1983 also, using flowering records which were available that year for many of the trees sampled. To first determine if such a stratification held promise, trees within each orchard were placed into early, mid and late flowering groups. Estimates of CE, SE, and FS were calculated for each group and tested simultaneously for equivalence (i.e., $\mu_{\text{early}} = \mu_{\text{mid}} = \mu_{\text{late}}$). The benefit of stratification was evaluated only when differences among phenological group means were detected ($p < 0.05$), with the intent of considering other cases if worthwhile benefits were found.

It was not possible to evaluate alternative subsampling allocations at each of the three or four stages, because only a single whorl and then a branch were selected per tree. However, evaluation of selecting one versus two branches per tree for CE estimation was feasible from the 1983 data because two branches were independently selected on several (11-23) trees at each orchard. The sampling of multiple cones per branch in 1984 and 1985 allowed a similar appraisal of sampling one, two, or three cones per branch for estimation of SE.

Absolute errors, with a 90% confidence interval, of ± 0.1 for CE and SE and ± 5 seeds for FS were employed to calculate sample size requirements. A rate of \$11/h, the wage of junior staff in these orchards in 1983-1985, was used to calculate the implementation costs for each approach.

Results

The CE, SE, and FS estimates varied considerably within each orchard over the years evaluated (Table 5). Phenology group differences were detected only for CE at Dewdney and Koksilah in 1983. Early flowering trees had lower CE values than mid or late trees at both orchards. At Koksilah, where the differences were most pronounced, CE values were 0.40, 0.76, and 0.81 for early, mid, and late flowering groups, respectively ($F = 17.41$; $df = 2,72$; $p < 0.01$).

TABLE 6. Number of trees required to estimate CE by means of three-stage sampling, and SE or FS by means of four-stage sampling (90% confidence)

Orchard	Year	Absolute sampling error			Sampling error relative to mean ($\pm 20\%$)		
		CE (± 0.10)	SE (± 0.10)	FS ^a (± 5 seeds)	CE	SE	FS ^a
Dewdney	1983	31	16	39 (14)	22	13	17 (7)
	1984	10	8	23 (9)	220	28	42 (15)
	1985	27	13	42 (15)	31	31	47 (15)
Koksilah	1983	28	12	33 (12)	16	18	29 (9)
	1984	37	4	10 (4)	78	176	253 (69)
	1985	23	14	29 (11)	12	22	34 (10)
Quinsam	1983	15	15	42 (15)	7	11	19 (7)
	1984	24	10	19 (7)	120	76	117 (39)
	1985	35	11	30 (11)	33	15	28 (10)
Snowdon	1983	25	17	37 (13)	13	20	32 (11)
	1984	22	6	8 (3)	195	110	142 (38)
	1985	21	20	32 (12)	124	22	37 (11)

NOTE: Allocation of a single branch and cone per tree for CE, and SE or FS, respectively.
^aThe number in parentheses equals the optimal number of cones dissected.

Stratification was therefore appraised for this variable and these orchards.

For CE estimation, D values for the initial three-stage method implemented ranged from 3.4 to 7.0, relative to the four-stage alternative method. Several times more conelets would therefore be required to achieve an estimate of equal precision if three-stage sampling were used. The associated number of trees required for the three-stage method was, however, consistently less for each orchard and year evaluated because these conelets were grouped on branches ($n_1 = 25$, on average, versus 51). Implementation costs were also always lower for the three-stage method because the added cost of assessing the greater number of conelets and cones was more than offset by the saving in the cost of within-orchard travel associated with the lower numbers of trees required. The implementation costs averaged \$99 over all orchards and years for the three-stage method versus \$144 for the four-stage method.

Stratified four-stage sampling produced only marginal improvements over four-stage sampling in the two orchards evaluated in 1983. At Dewdney, the implementation cost was \$179, while at Koksilah it was \$141 for stratified four-stage sampling compared with unstratified sampling costs of \$183 at Dewdney and \$163 at Koksilah. Implementation costs for this alternate method would therefore also be greater than those of the three-stage sampling method implemented initially.

Stratified three-stage sampling did yield considerable improvement over three-stage sampling, but only at Koksilah, the orchard with the largest among-strata differences. Sample size was reduced from 28 to 18 trees, which would have resulted in an implementation cost reduction of \$41. At Dewdney, n_1 was reduced by only two trees, with an associated cost saving of \$9.

Allocation of two branches per tree exhibited no consistent advantage over single branch sampling. The implementation cost was less at Quinsam (\$65 vs. \$79), greater at Snowdon (\$158 vs. \$112), and comparable at Dewdney (\$131 vs. \$135) and Koksilah (\$109 vs. \$115).

For SE, D values for four-stage sampling ranged from 7.9 to 18.2 and were always much higher than those for pro-

cessing individual fertile sites in five-stage sampling for SE ($D = 1.0$). However, the implementation costs required to estimate SE, using the four-stage method implemented, were consistently less because many fewer trees were required. The numbers of trees and implementation costs averaged 12 and \$57, respectively, for four-stage sampling, whereas for five-stage sampling these values averaged 74 and \$169, respectively. Allocation of one cone per branch with four-stage sampling consistently exhibited the lowest implementation costs, averaging \$49 (over 1984–1985) compared with \$73 for two cones per branch and \$93 for three cones per branch.

For FS, D values for two-phase sampling averaged 1.3, and were consistently higher than those for the single-phase approach ($D = 1.0$). The optimal proportion of dissected to total cones sampled (P_d) ranged from 0.26 to 0.35. Associated sample size requirements were larger, averaging 28 trees compared with 21 trees for single-phase sampling. The implementation costs were, however, consistently lower for two-phase sampling, averaging \$81 compared with \$102, because only about one-third of the cones needed to be fully dissected. Plots of the full cone counts (Y_i) on the slice counts (X_i) revealed a linear relationship, and FS estimates derived using eq. 2 should therefore not be seriously biased.

Subsequent efficiency estimation

Based on these results it was recommended that the three-stage method, with allocation of a single branch per tree, be generally adopted for subsequent CE estimation in coastal Douglas-fir orchards. Stratification by phenology group did yield a lower implementation cost, but this was dependent upon low within-strata variation which did not always materialize. Sample selection would also be more difficult because separate within-strata samplings would be required. The appropriate mean and variance estimators would also be more complex. These drawbacks outweighed the cost savings demonstrated.

For SE, the four-stage method, with allocation of a single cone per tree, was consistently least costly to implement and was therefore recommended. For FS, the alternative four-stage method with two-phase sampling of cones at the last

stage was prescribed, again with allocation of a single cone per tree. In general, about one-third of the cones sampled should be fully dissected.

Sample size requirements (n_1 values) for these recommended methods and allocations are given in Table 6 for each orchard and year evaluated. Absolute errors of ± 0.1 for CE and SE and ± 5 seeds for FS, which is consistent with survey objectives, and an arbitrary relative error of $\pm 20\%$ were both used. Requirements for the absolute errors for each variable were of the same magnitude and it was feasible to make recommendations for subsequent surveys. Specifically, 40, 20, and 45 trees were suggested for CE, SE, and FS, respectively. Sampling simulations showed that such sizes were sufficient to yield near-normal sample mean distributions, thus allowing the calculation of valid confidence intervals. These sizes represent a cost of less than \$200 to estimate CE, and SE or FS for each orchard annually.

Requirements corresponding to the 20% error fluctuated sharply (Table 6), principally because of the erratic efficiency means experienced. Thus, it was not possible to prescribe future sample size requirements. Rather, pilot studies were recommended to first gauge n_1 before completing each survey. Unfortunately, because CE could not be estimated until cone harvest, n_1 for the initial branch counts would need to be large, i.e., about 200 branches, and subsequently reduced when appropriate as the final cone tallies were made and preliminary CE estimates derived.

Acknowledgments

We thank M. Senecal, P. Adamson, M. Curran, T. Crowder, G. Morrow, and D. Rudolph for technical assistance, D. Errico, M. Crown, C. Leadem, J. Webber, V. Fletcher,

C. Simmons, and H. Stauffer for advice, and L. Safranyik, M. Kovats, S. Omule, T. Shore, and anonymous reviewers for reviewing the initial drafts of the manuscript.

- ANONYMOUS. 1974. Seeds of woody plants of the United States. USDA For. Serv. Agric. Handb. No. 450.
- BRAMLETT, D.L., and GODBEE, J.F., JR. 1982. Inventory-monitoring systems for southern pine seed orchards. Ga. For. Comm. Res. Pap. No. 28.
- BRAMLETT, D.L., BELCHER, E.W., JR., DEBARR, G.L., HERTEL, G.D., KARRFALT, R.P., LANTZ, C.W., MILLER, T., WARE, K.D., and YATES, H.O., III. 1977. Cone analysis of southern pines: a guidebook. USDA For. Serv. Gen. Tech. Rep. SE-13.
- COCHRAN, W. 1977. Sampling techniques. 3rd ed. John Wiley and Sons, Toronto.
- DOBBS, R.C., EDWARDS, D.G.W., KONISHI, J., and WALLINGER, D. 1976. Guidelines to collecting cones in B.C. conifers. Joint Rep. No. 3, B.C. Forestry Service and Canadian Forestry Service, Victoria, B.C.
- SILEN, R.R., and KEANE, G. 1969. Cooling a Douglas-fir seed orchard to avoid pollen contamination. USDA For. Serv. Res. Note PNW-101.
- SUKHATME, P.V., and SUKHATME, B.V. 1970. Sampling theory of surveys with applications. 2nd ed. Food and Agriculture Organization, Rome.
- WHITE, T.L., HARRIS, H.G., JR., and KELLISON, R.C. 1977. Conelet abortion in longleaf pine. Can. J. For. Res. 7: 378-382.
- WINJUM, J.K., and JOHNSON, N.E. 1960. A modified knife-cone cutter for Douglas-fir seed studies. J. For. 58: 487-488.
- ZASADA, J.C. 1971. Frost damage to white spruce cones in interior Alaska. USDA For. Serv. Res. Note PNW-149.