Adaptive Cluster Sampling for Forest Inventories

FRANCIS A. ROESCH, JR.

ABSTRACT.   Adaptive cluster sampling is shown to be a viable alternative for sampling forests when there are rare characteristics of the forest trees which are of interest and occur on clustered trees. The ideas of recent work in Thompson (1990) have been extended to the case in which the initial sample is selected with unequal probabilities. An example is given in which the initial sample of trees is selected with probability proportional to tree basal area. If a characteristic of interest is observed on a sample tree, additional trees within a fixed distance of the sample tree are also included in the sample. FOR. SCI. 39(4):655–669.
ADDITIONAL KEY WORDS.   Forest health, biodiversity, sequential sampling.

FOREST SURVEYS ARE OFTEN CONCERNED WITH the simultaneous evaluation of many characteristics of the forest. In addition, many of these characteristics occur only rarely, especially in extensive surveys such as those conducted by the USDA Forest Service's Forest Inventory and Analysis Units (FIA). Many forest surveys were designed with the idea that wood was the primary characteristic of interest in the forest. These surveys were therefore intended to sample wood efficiently. While wood is an important benefit of the forest, we want to sample other characteristics of the forest efficiently as well. The results in this paper are especially attractive because it is not always possible or desirable to establish individual sample designs for each characteristic of interest. We show how a general framework for adaptive cluster sampling could be used in a forest survey design to efficiently sample many items of interest simultaneously.

Adaptive cluster sampling can be a highly efficient way to sample clustered rare events. First, a probability sample of units in a population is taken and then additional units are selected near (or in the neighborhoods of) those units that display any of the rare events of interest. Any additional observations of the same rare event will trigger further sampling in the neighborhoods of the new units possessing the rare event. This cycle is repeated until no new units display the rare event.

Although adaptive sampling ideas have been around for quite some time [at least since Wald (1947)], this presentation is the first to combine the probability proportional to size sampling schemes common in forestry with an adaptive sampling scheme. The result is a system which could be applied immediately, at a relatively small cost, to many already in-place forest inventory systems. A thorough introduction to adaptive designs for the case in which the initial sample is selected with equal probability can be found in Thompson (1990), which provided some of the theoretical foundation for the development below. The interested

reader might also examine the related work in Birnbaum and Sirken (1965), Levy (1977), Thompson (1991a), and Thompson (1991b). The general topic of sampling rare populations is discussed in Kalton and Anderson (1986).

We will examine the simple case of a single rare characteristic of interest. Suppose that the tree is the sampling unit and that there are N trees in the forest with labels $1, 2, \ldots , N$. Associated with the N trees are values of interest $y = \{y_1, y_2, \ldots , y_N\}$. We are interested in estimating the population total of the y-values (T).

For every tree i in the population, we define a neighborhood A_i consisting of a collection of trees which includes tree i. Assume these neighborhoods group trees which are physically close to each other. Also, the neighborhoods are symmetric: if tree i is in the neighborhood of tree j, then tree j is also in tree i's neighborhood.

Suppose this rare characteristic of interest is the presence of a certain type of pollution damage, which we will call pollution damage D, on the trees in our forest. If tree i exhibits pollution damage D, set $y_i = 1$, otherwise set $y_i = 0$.

We choose the initial sample of trees in the forest by first selecting sample trees from a randomly located point in the forest by some rule. Two common selection rules are those corresponding to fixed-area plot sampling and horizontal point sampling. Without loss of generality, we will assume that the initial selection of trees is by horizontal point sampling, which is a method of selecting sample trees with probability proportional to the cross-sectional area of the tree bole at 4.5 ft above the ground (basal area). Each tree is surrounded by a circular area of selection, the size of which depends on basal area, and from within which a randomly placed point will select it for the sample. Readers unfamiliar with this sampling technique are referred to Husch et al. (1982), Avery and Burkhart (1983), or Grosenbaugh (1958). If an initial sample tree j has pollution damage D (i.e., $y_j = 1$), then we could use tree j's center as the center of a circle of radius r. We observe the value of y for all of the other trees within this circle which have not already been selected for the sample from this point. If, for any of those trees, $y_i = 1$, we do the same thing we did for the first set of y_j's equal to 1. We stop when we find no new trees for which y_i equals 1.

A cluster j is the set of all trees included in the sample as a result of the initial selection of tree j. Within a cluster, a network is defined as the subset of trees such that selection of any tree within the network will lead to the inclusion of every other tree in the network. Any tree not possessing pollution damage D, but in the neighborhood of one that does is termed an *edge* tree. Figure 1 illustrates these definitions. In Figure 1, tree 6 is an edge tree with respect to both trees 3 and 4, while tree 5 is an edge tree with respect to tree 4. Inclusion of an edge tree will not result in the inclusion of any other trees. If we consider all trees not possessing pollution damage D as networks of size 1, we have uniquely partitioned the tree population into M networks, conditioned on pollution damage D. We can do this for many kinds of rare events, in which case the tree population can be visualized as many separate overlays, each containing a population of networks of a different characteristic.

Each tree in the initial sample, i.e., each tree selected by the randomly placed point, is sampled with known probability. In the estimators that follow, we will use only those trees that were in the initial sample and those additional trees for which $y_i = 1$. This allows us to calculate the probability of using a sample tree in the estimator of a given characteristic, even though we do not (and cannot) know its
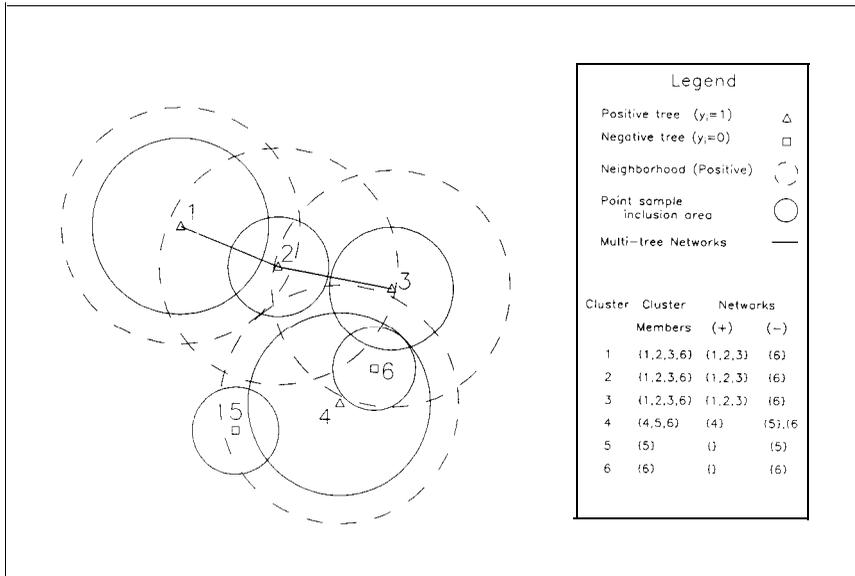
FIGURE 1.   Adaptive sampling attributes for a group of six trees in a population. One randomly placed point can select the following sets of trees for the initial sample: { }, {1}, {3}, {2}, {4}, {5}, {1,2}, {3,4}, {3,4,6}, {4,5}, and {4,6}. For clarity, the neighborhoods for negative trees are not shown.

probability of inclusion in the sample. To see this, consider the probability of tree *i* being selected for the sample from any individual random point:

$$p_i = p_I = \frac{u_I \cup a_I}{L} \tag{1}$$

where

a_I = union of the selection areas for the trees in network I to which tree i belongs,

u_I = union of the selection areas for all the trees in networks of which tree i is an edge unit,

L = the total area of the forest.

(Note that capitalized subscripts are used when the quantity pertains to the network and lowercase subscripts are used when the quantity pertains to the tree. In the sequel, when both a lowercase and uppercase version of the same subscript appear in an equation, it indicates a mapping of the tree population into the network population.) This sampling scheme does not provide knowledge of p_i for all trees since u_I is not **known** when $y_i = 0$, but it does allow knowledge Of p_i for trees possing the characteristic of interest (for these trees $u_I = 0$). We know the probability for each tree used in the estimator, because trees for which $y_i = 0$ selected in the adaptive part of the sample are not used in the estimator.

ESTIMATORS OF T

The following estimators assume that we randomly locate *m* points on the surface of a forest and that we are sampling with replacement. The total of the observations over network K is

$$y_K = \sum_{j=1}^{v_K} y_j \tag{2}$$

where v_K is the number of trees in network K. The mean per unit area for network K is

$$\bar{y}_K = \frac{y_K}{a_K} \tag{3}$$

Thompson (1990) presents a modified Hansen-Hurwitz (Hansen and Hurwitz 1943) type unbiased estimator of the total character of interest, which we will modify again to allow a random number of networks (M_h) sampled from each point h:

$$t_{HH} = \left(\frac{L}{m}\right) \sum_{h=1}^{m} \sum_{K=1}^{M_h} U_K \tag{4}$$

Note that t_{HH} counts each network once for each time it is selected by a randomly placed point, no matter how many of its component trees are selected at that point. The probability of using tree i in t_{HH} is proportional to the union of the selection areas of the trees in the network to which tree i belongs divided by the area of the forest. After a derivation similar to that found in Roesch (1990), the variance of t_{HH} can be shown to be equal to

$$\mathrm{var}(t_{HH}) = \left(\frac{1}{m}\right) \sum_{I=1}^{M} \sum_{J=1}^{M} (O_{IJ}L - a_I a_J)\bar{y}_I \bar{y}_J \tag{5}$$

where O_{IJ} is the area of overlap of the selection areas of networks I and J. An estimator of the variance is:

$$s^2(t_{HH}) = \left[\frac{1}{m(m-1)}\right] \sum_{h=1}^{m} \left(\left(L \sum_{K=1}^{M_h} \bar{y}_K\right) - t_{HH}\right)^2 \tag{6}$$

A modification of t_{HH} would result in another unbiased estimator if we count a network once for each of its component trees selected by the initial point sample. That is, if three of a network's trees are selected at a point, then the network is counted three times and the probability of selecting the network becomes proportional to the sum of the component trees' probabilities rather than their union. For each tree k selected by the initial point sample, the sum of the y-values in the tree's network K divided by the sum of the areas of selection for each of these trees is:

$$\tilde{y}_k = \frac{y_K}{\displaystyle\sum_{j=1}^{v_K} a_j} \tag{7}$$

where

$$a_j = \text{the selection area of tree j.}$$

This leads to the estimator:

$$t_{HHM} = \frac{L}{m} \sum_{h=1}^{m} \sum_{k=1}^{n_h} \tilde{y}_k \tag{8}$$

where n_h is the number of trees selected by random point h.
The variance of t_{HHM} is:

$$\text{var}(t_{HHM}) = \left(\frac{1}{m}\right) \sum_{i=1}^{N} \sum_{j=1}^{N} (o_{ij}L - a_i a_j)\tilde{y}_i\tilde{y}_j \tag{9}$$

where

$o_{ij} = $ the area of overlap of the selection areas of trees i and j.

An estimator of the variance is:

$$s^2(t_{HHM}) = \left[\frac{1}{m(m-1)}\right] \sum_{h=1}^{m} \left(\left(L \sum_{k=1}^{n_h} \tilde{y}_k\right) - t_{HHM}\right)^2 \tag{10}$$

Thompson (1990) shows that an unbiased estimator can be formed by modiig the Horvitz-Thompson estimator (Horvitz and Thompson 1952) to use observations not satisfying the condition only when they are part of the initial sample. The probability of tree k, in network K, being included in the sample from at least one of the m random points is:

$$\alpha_k = \alpha_K = 1 - \left(1 - \frac{a_K}{L}\right)^m \tag{11}$$

A modified Horvitz-Thompson type estimator is

$$t_{HT} = \sum_{k=1}^{n} \left(\frac{y_k J_k}{\alpha_k}\right) \tag{12}$$

where:

$$J_k = \begin{cases} 0 \text{ if the kth tree does not satisfy the condition and is not included in} \\ \quad \text{the initial sample} \\ 1 \text{ otherwise} \end{cases}$$

and n is the total number of distinct trees samples. For convenience, we will work with the networks in the variance formulas. The joint probability of inclusion of networks J and H at least once from the m random points is equal to:

$$\alpha_{JH} = 1 - \left\{\left[1 - \frac{a_J}{L}\right]^m + \left[1 - f\right]^m - \left[1 - \frac{U_{JH}}{L}\right]^m\right\} \tag{13}$$

where U_{JH} is the union of the selection areas of networks J and H. The variance of t_{HT} is

$$\text{var}(t_{HT}) = \sum_{J=1}^{M} \sum_{H=1}^{M} y_J y_H \left(\frac{\alpha_{JH} - \alpha_J \alpha_H}{\alpha_J \alpha_H} \right) \tag{14}$$

and an estimator of the variance is:

$$s^2(t_{HT}) = \sum_{K=1}^{\kappa} \sum_{H=1}^{\kappa} y_K y_H \left(\frac{\alpha_{KH} - \alpha_K \alpha_H}{\alpha_K \alpha_H \alpha_{KH}} \right) \tag{15}$$

where κ equals the number of distinct networks in the initial probability sample.

METHODS

DATA DESCRIPTION

The data were a subset of those collected by the USDA Forest Service Northeastern Forest Experiment Station in Hancock County, Maine. Fifty-three circular, concentric $\frac{1}{10}$ ac pulpwood plots and $\frac{1}{5}$ ac sawlog plots were established in 1968 and remeasured in 1981. All trees that were at least 5 in. in diameter at breast height (dbh) were measured on the inner $\frac{1}{10}$ ac, while only trees of sawlog size (at least 9 in. dbh for softwoods and at least 11 in. dbh for hardwoods) were measured on the outer $\frac{1}{10}$ ac. The azimuth and distance from plot center for each tree were recorded to the nearest degree and $\frac{1}{10}$ foot respectively. The only tree information used here is species, dbh, and location for 1981.

SIMULATION

One issue that is particularly difficult to address in extensive forest surveys is that of biodiversity, simply because very imprecise estimates of a characteristic are often obtained for rare species. In the following example, the data described above are used in a computer simulation to test the effectiveness of the point sampling/adaptive cluster sampling scheme for estimating the number of trees of certain species displaying varying levels of rarity and aggregation. The simulation was written in GAUSS386 Version 2.2 and run on a DELL 486/50 MHZ personal computer.

For the simulation, a highly diverse "forest" was created by cutting the largest square possible out of the inner $\frac{1}{10}$ ac plots, each side of the square facing one of the cardinal directions, and then using the first 49 of these squares (by plot number) in a 7 x 7 arrangement to simulate a square forest of approximately 3.11 ac. Figure 2 displays the location of each tree in the forest and then the individual locations of the eight "rare" species, which are arranged roughly in order of increasing aggregation. The identities of the species have been retained to facilitate the discussion rather than for any particular significance of the species themselves. What is important is the spatial relationship between the members of
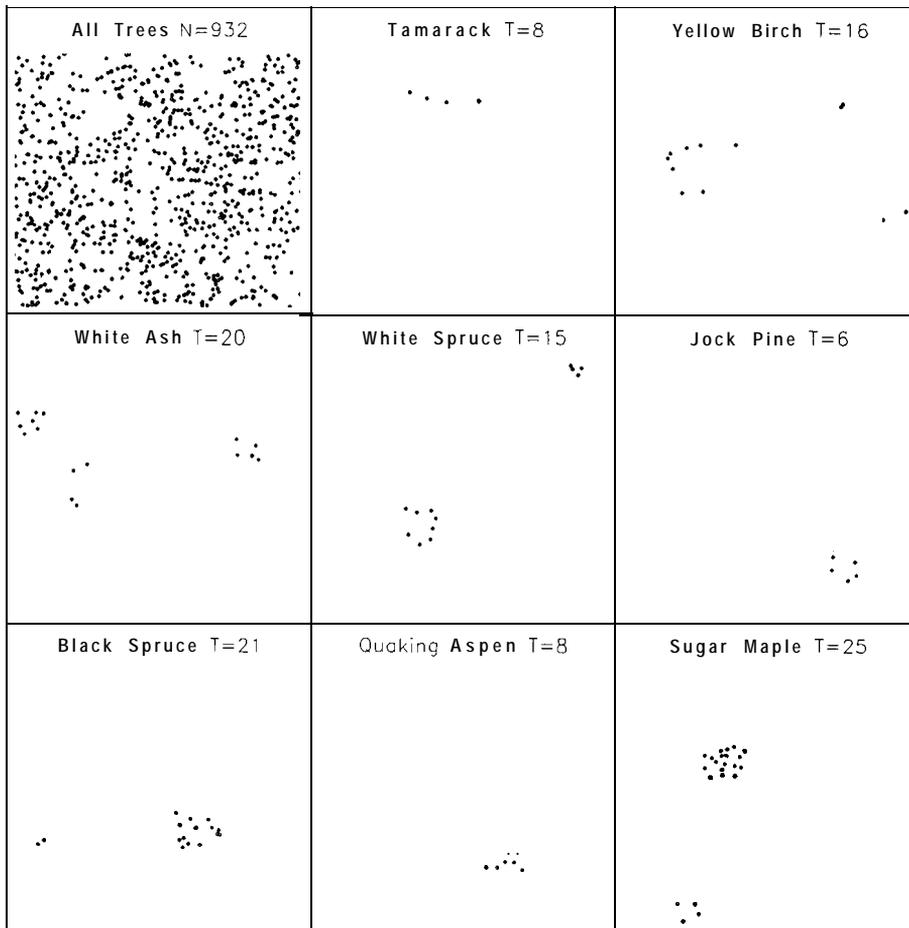
FIGURE 2. The spatial locations on the 3.11 ac simulated forest of all trees (upper left) and of members of each of the species of interest.

each of the species of interest, e.g., those relationships found in Figures 2 through 4.

For the original probability sample, ten random points were located in the forest, and a Basal Area Factor (BAF) of 10 ft^2/ac was used. This sample design was then replicated 1000 times in the simulation. Figure 3 depicts the point sample area of selection for each member tree of the eight rare tree species. For the adaptive part of each sample, the inverse of the search area in acres was varied from 10 to 90 by 20; e.g., search areas of 1/10, 1/30, 1/50, 1/70, and 1/90 of an acre were used. When a member of a species of interest was found by the point sample, the search area centered at that tree was checked for more trees of that species. As described above, any new trees found of that species became the centers of search circles until all the trees in the cluster were identified. The additional cost of including these extra trees in the sample consists of identifying the species, and if the tree is one of the rare species, measuring the dbh and recording the location of the tree. If this species is truly rare and found in clusters, then the additional cost should be small, because of the rareness of the species,

1/10	1/30	1/50	1/70	1/90	Search Area (acres) / Species
5	8	8	8	8	Tamarack
7	12	14	.14	14	Yellow Birch
5 / 7	6	10	11	14	White Ash
3 / 3	3 / 4	6	10	10	White Spruce
1	2	4	4	4	Jock Pine
6	8	9	12	12	Black Spruce
2	2	2	3	6	Quaking Aspen
3	5	5	9	11	Sugar Maple

FIGURE 3. The point-sample selection areas for the members of the eight rare species for a BAF of 10. A tree is selected for the sample if a point lands within its selection area.

and the estimate of the number of trees of that species may be improved, because of the clustering of the species.

Changing the search area changes the size of the clusters found for each "positive" tree in the point sample. Figure 4 shows the positive networks for each of the "rare" species at each search area size. For each of the eight relatively rare

FIGURE 4. The positive networks formed for each species at each search area size. The numeral in each plot is the number of positive networks.

species, at each search circle size, both the mean and the mean squared error (MSE) were calculated for four estimators of the number of trees. Three of the estimators are those utilizing the additional adaptive cluster sample, and defined above, (t_{HH}, t_{HHM}, and t_{HT}) and the fourth is the normal point sample estimator:

$$t_G = \frac{L}{m} \sum_{h=1}^{m} \sum_{i=1}^{n_h} \frac{y_i}{a_i} \tag{16}$$

which is included as a check. The extra work involved in the adaptive sampling scheme is not justified if the resulting estimate would not at least be better than the normal point sample estimate. In addition, the estimated variance was calculated for each of the three adaptive sampling estimators.

RESULTS

Figure 5 shows the relative bias [$(\hat{E}[t_i] - T)/T$], where $\hat{E}[t_i]$ is the simulation mean for estimator t_i; $i = G, HH, HHM$, and HT] for the four estimators versus
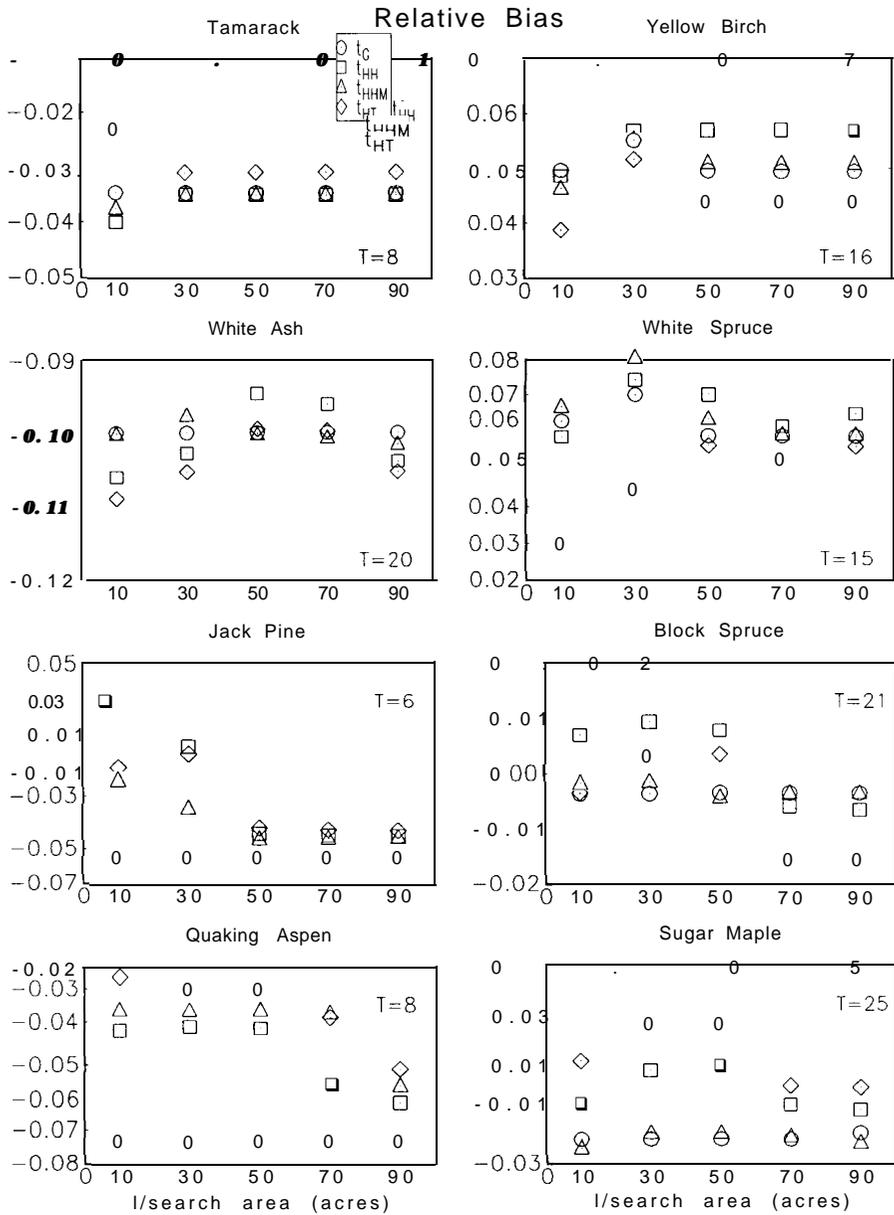
FIGURE 5. The relative bias of the four estimators by inverse of the search circle size in acres, from 1000 simulations of 10 randomly placed points, for each of the 8 rare species.

the inverse of the search area, obtained from the 1000 sets of estimates, based on ten randomly placed points each. Overall, Figure 5 confirms the theoretical unbiasedness of the adaptive sampling estimators, as there is no evidence of additional bias due to the adaptive sampling technique. What little apparent bias there is occurs in all of the estimators, including t_G, which does not rely on the adaptive part of the sample.

An additional statistic useful for comparison is MSE, which equals the variance for an unbiased estimator. Figure 6 displays the ratio of MSE for each of the three
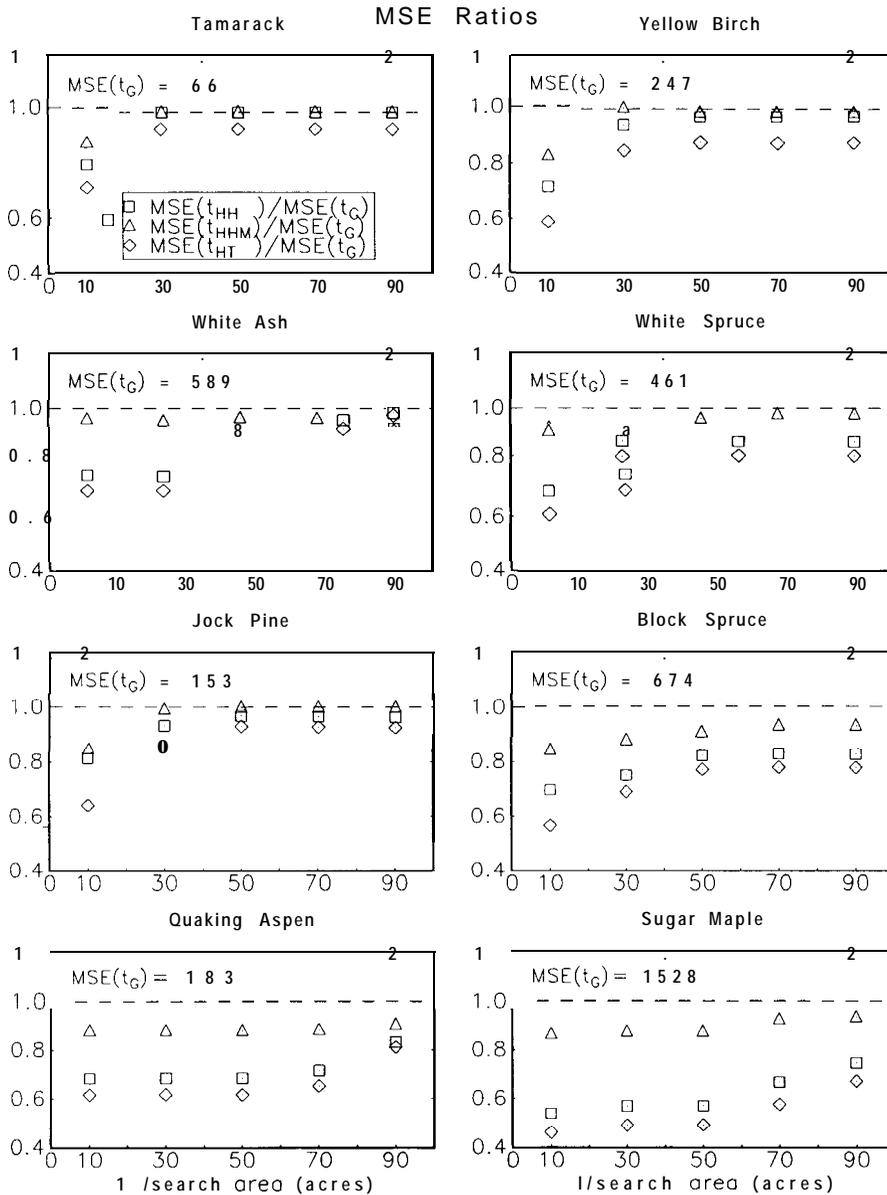
Tamarack

MSE(t_G) = 66

□ MSE(t_{HH})/MSE(t_G)
△ MSE(t_{HHM})/MSE(t_G)
◇ MSE(t_{HT})/MSE(t_G)

Yellow Birch

MSE(t_G) = 247

White Ash

MSE(t_G) = 589

White Spruce

MSE(t_G) = 461

Jock Pine

MSE(t_G) = 153

Block Spruce

MSE(t_G) = 674

Quaking Aspen

MSE(t_G) = 183

1 /search area (acres)

Sugar Maple

MSE(t_G) = 1528

l/search area (acres)

FIGURE 6. The ratio of the mean squared error for each of the three adaptive sampling estimators to the mean squared error of the point sampling estimator, by the inverse of the search circle size, in acres, from 1000 simulations of 10 randomly placed points, for each of the 8 rare species.

adaptive sampling estimators to the MSE of t_G versus the inverse of the search area, again obtained from the 1000 sets of 10 randomly placed points. We can see the advantage of the adaptive strategy by examining the MSE ratios in this figure. Also evident is the apparent superiority of t_{HT} over t_{HH} and t_{HHM} for use with this adaptive sampling strategy. The MSE of t_{HT} is never greater than the MSE of t_G. In general, the larger the search area, the smaller the MSE of the estimators using the adaptive design, once the thresholds defining the network sizes are

exceeded. For the less clustered species, there is some advantage to the adaptive strategy if the search area is large enough. For example, in the graphs for tamarack and yellow birch, a substantial reduction in MSE is apparent for t_{HH}, t_{HHM}, and t_{HT} when the search area is $\frac{1}{10}$ ac. For tamarack, Figure 6 shows no MSE advantage to the adaptive strategy in the smaller search areas if t_{HH} and t_{HHM} *are* used, while a slight lowering of MSE is shown for t_{HT}. These results are consistent with the observation that the tamarack trees would not appear together in networks until the search circles are rather large. For example, referring to Figure 4, for the four smallest search circles, the networks of tamarack each contained one tree. The largest search circle of $\frac{1}{10}$ ac resulted in five networks of tamaracks, four networks containing only one tree and one network containing four trees. This illustrates that the adaptive strategy will only be advantageous for traits found on clustered trees. From the upper lefthand plot in Figure 2, depicting the positions of all trees, the reader will note that quite a few trees of other species had to be examined each time a tamarack tree was found. For this species, this resulted in increased cost with a benefit of only slightly reduced MSE until the largest search circle was used.

Examining the results for white ash and white spruce in Figure 6, we see roughly two groupings Of the MSEs (t_{HT} with t_{HH} and t_G with t_{HHM}), with an advantage to t_{HT}. For the species in which clustering is pronounced, white ash, white spruce, black spruce, quaking aspen, and sugar maple, Figure 6 shows t_{HT} to be always best, followed by t_{HH}, t_{HHM} and finally t_G. The jack pine results in Figure 6 show a distinct advantage to the adaptive strategy only at the two largest search circles.

For each Of the 1000 iterations Of the simulation, $s^2(t_{HH})$, $s^2(t_{HHM})$, and $s^2(t_{HT})$ were calculated for each search circle size. Figure 7 shows the ratio of the mean of these estimates to the respective MSEs for the estimators of the total. This ratio should be very close to 1 for an unbiased estimator of the variance of an unbiased estimator of the total. We see that this is indeed the case with ratios always falling in the interval $(0.9, 1.1)$ and usually falling in the tighter interval $(0.95, 1.05)$.

DISCUSSION

In general, it appears from this simulation that when an adaptive strategy is called for, t_{HT} can be expected to perform well in terms of MSE. It performed best by that criterion here. This is the same result obtained by Thompson (1990) for the equal probability case. Occasionally, t_{HH} is approximately equivalent to t_{HT} by the MSE criterion. Thompson (1990) points out that there should be populations for which the Hansen-Hurwitz type estimators are superior; however, the conditions leading to this superiority were not revealed in either that work or this work. This simulation did not uncover any advantage to using t_{HHM}. When the MSE of t_{HHM} is lower than that of t_G, it is still not as low as the MSE of the other two adaptive sampling estimators. Reductions in MSE of over 50% have been demonstrated for t_{HT} over t_G in some cases. This size of a reduction in MSE would often be critical when sampling rare, clustered events such as rare tree species, insect infestations, and certain diseases.

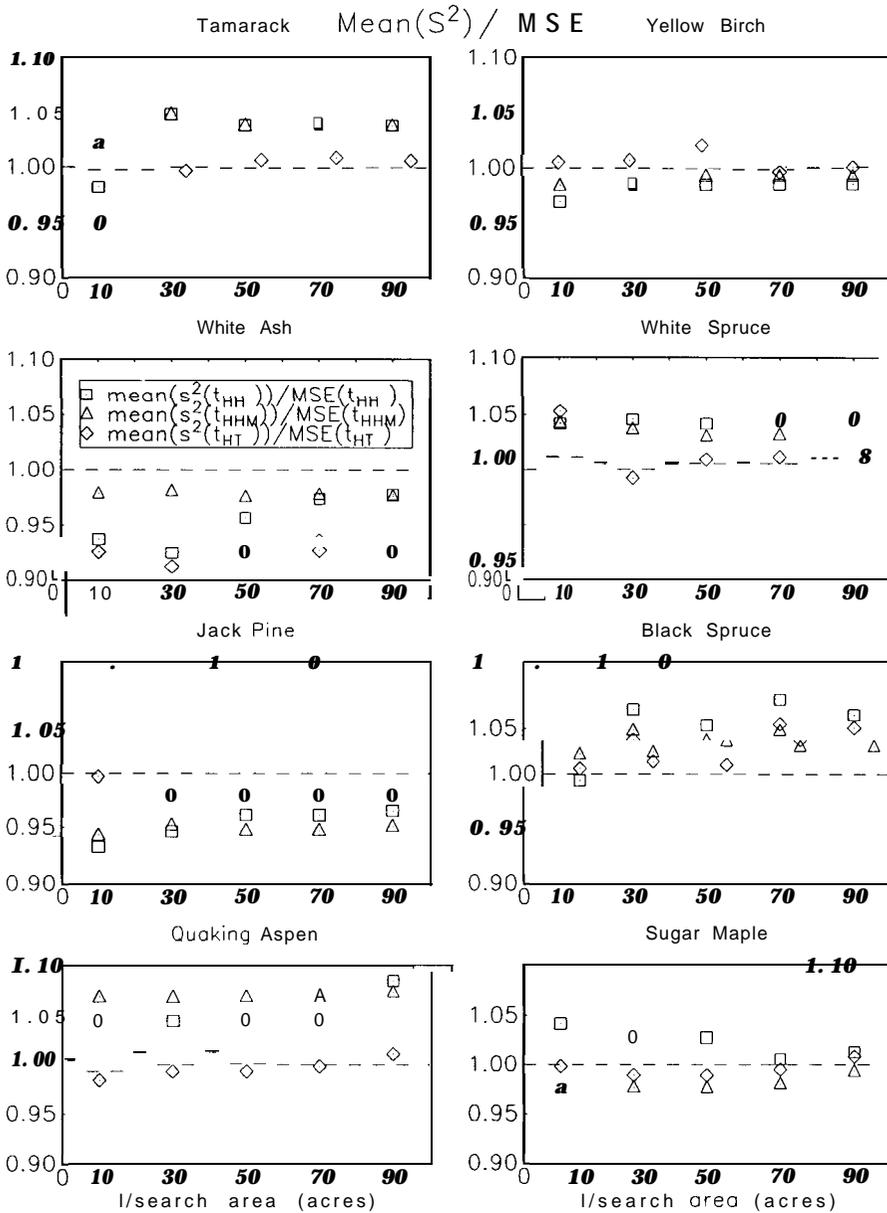A major concern when choosing between sampling strategies is the relative

FIGURE 7. The ratio of the mean estimate of variance to the mean squared error for each of the adaptive sampling estimators, plotted against the inverse of the search circle size, in acres, from 1000 simulations of 10 randomly placed points, for each of the 8 rare species.

cost of the strategies. Cost is evaluated in this paper as the size of the search circle necessary for a given improvement in MSE. The additional monetary cost of the adaptive strategy for a particular application depends on relative cluster size and occurrence in the sample. These factors can be controlled by the inventory designer, given adequate prior knowledge of the population.

Due to the additional cost, the adaptive strategy should be saved for characteristics that are considered very important, and known to be both rare and found

in clusters. The requirement that the trait motivating the adaptive strategy be rare is one of particality, not one of necessity from a statistical point of view. Even when an extremely important characteristic is known to be both rare and clustered, it may be more appropriate to design an individual survey to estimate the characteristic, rather than to force the estimation into an existing inventory. Given the decision to make the estimate from an existing inventory, the estimate of the marginal cost of adapting for the characteristic must be weighed against the benefit of a better estimate.

There are two major advantages of the adaptive sampling strategy. The first is that the decision to use an adaptive scheme can be made on a characteristic by characteristic basis so the decision to adapt for one characteristic does not affect the cost of estimating other characteristics. The second advantage is that only the presence of the characteristic triggers additional cost. Most other solutions to the rare event problem, such as increasing the size or number of plots, do not have these two properties.

A problem which was not addressed in the simulation is the case-specific amount of risk involved in misidentifying the characteristic of interest. With a nonadaptive sample, one could hope that misidentifications would compensate for each other. With an adaptive sample, this hope would often be unrealistic. Therefore, in addition to the conditions given above, an adaptive sampling scheme should be applied only to those characteristics that can be identified with complete confidence.

Although we have used a fixed radius from tree center to choose the additional sample, many other measures of closeness could be conceived. Whatever measure is chosen, it should be scaled to define clusters which are as large as practicable.

In this paper, I have extended the theory for adaptive cluster sampling to the case when the initial sample units are chosen with probability proportional to size, as is often the case in forest inventories. This will allow an easy integration of an adaptive scheme into many already existing forest inventories. This ability will become more important in the future as forest inventory specialists are expected to answer a more diverse set of questions.

LITERATURE CITED

AVERY, T.E., and H.E. BURKHART. 1983. Forest measurements. McGraw-Hill, New York. 331 p.

BIRNBAUM, Z.W., and M.G. SIRKEN. 1965. Design of sample surveys to estimate the prevalence of rare diseases: Three unbiased estimates. Public Health Serv. Publ. No. 1000-Series 2-No. 11. 8 p.

GROSENBAUGH, L.R. 1958. Point-sampling and line-sampling: Probability theory, geometric implications, synthesis. USDA For. Serv. Occas. Pap. 160. 34 p.

HANSEN, M.H., and W.N. HURWITZ. 1943. On the theory of sampling from finite populations. Annals Math. Stat. 14:333–362.

HORVITZ, D.G., and D.J. THOMPSON. 1952. A generalization of sampling without replacement from a finite universe. JASA 47:663–685.

HUSCH, B., MILLER, C.I., and T.W. BEERS. 1982. Forest mensuration. Wiley, New York. 402 p.

KALTON, G., and D.W. ANDERSON. 1986. Sampling rare populations. J.R. Stat. Soc. A 149:65–82.

LEVY, P.S. 1977. Optimum allocation in stratitied random network sampling for estimating the prevalence of attributes in rare populations. JASA 72:758–763.

ROESCH, F.A., JR. 1990. The variances of some compatible estimators for the components of growth from remeasured point samples. P. 189-196 in State-of-the-art methodology of forest inventory: A symposium proceedings. USDA For. Serv. Gen. Tech. Rep. PNW-263.

THOMPSON, S.K. 1990. Adaptive cluster sampling. JASA 85:1050–1059.

THOMPSON, S.K. 1991a. Stratified adaptive cluster sampling. Biometrika 78:389–397.

THOMPSON, S.K. 1991b. Adaptive cluster sampling: Designs with primary and secondary units. Biometrics 47:1103–1115.

WALD, A. 1947. Sequential analysis. Wiley, New York.

AUTHORANDACKNOWLEDGMENTS