Techniques and Computations for Mapping Plot Clusters that Straddle Stand Boundaries

CHARLES T. SCOTT
WILLIAM A. BECHTOLD

ABSTRACT. Many regional (extensive) forest surveys use clusters of subplots or prism points to reduce survey costs. Two common methods of handling clusters that straddle stand boundaries entail: (1) moving all subplots into a single forest cover type, or (2) "averaging" data across multiple conditions without regard to the boundaries. These methods result in biased estimates of tree attributes (volume/acre, basal area/acre, etc.) and/or misclassification of discrete stand-area attributes (cover type, stand size, etc.). To circumvent the problems encountered with these approaches, a technique is described whereby subplots always are arranged in a fixed pattern, and trees are tallied by the cover type or "condition class" in which they occur. This involves mapping boundaries that bisect a subplot and proportioning subplot area by condition class. Field and calculation procedures leading to unbiased estimates are given. FOR. SCI. MONOGR. 31:46–61.
ADDITIONAL KEY WORDS: Forest inventory, cluster sampling.

TO THE EXTENT POSSIBLE, FOREST MENSURATIONISTS confine sampling efforts to forested conditions. Plots on nonforest land usually are ignored, but complications arise when sample plots straddle the edge between forest and nonforest, or cross the boundary between two distinctly different forest cover types. If plots straddling multiple conditions are handled improperly, significant bias and/or classification errors can result. Finney and Palca (1948) were among the first to recognize the problem and to provide a possible solution, biased though it was. Schmid-Haas (1969) and Gregoire and Scott (1990) present and evaluate a variety of boundary methods.

Several techniques have been proposed to overcome edge bias. These include the border-zone method (Grosenbaugh 1958), the expanded tree-circle method (Barrett 1965), the reflection method (Schmid-Haas 1969), and the tree-concentric method (Fowley and Arvanitis 1981, Gregoire and Scott 1990). Unfortunately, these methods are too difficult to be applied to cluster sampling (the preferred sampling method for extensive surveys), because they depend on a regularly shaped plot around a single point. The plot shape represented by a cluster of subplots is far too complicated for these methods. Exacerbating the problem, overlapping conditions occur more frequently with clusters because they cover a wider area.

The objectives of this study were to: (1) develop estimation procedures that result in unbiased plot-level and population-level estimates of tree attributes (numbers of trees/acre, volume/acre, basal area/acre, etc.) and stand characteristics (forest cover type, stand size, etc.); and (2) accomplish this through practical field procedures applicable to plot clusters in a consistent manner.

BACKGROUND

Unlike management-level forest inventories where stand boundaries are mapped and areas are known, extensive surveys must sample across entire regional land bases to estimate forest area. In the United States, regional surveys are conducted on a state-by-state basis by the six Forest Inventory and Analysis (FIA) units of the USDA Forest Service (USDA For. Serv. 1987). These same FIA units also participate in regional surveys conducted by the national Forest Health Monitoring (FHM) program (Scott et al. 1993). All of these surveys are based on some form of cluster sampling, primarily because it is the most cost-effective (Scott et al. 1983). The experiences of FIA and FHM with cluster sampling and edge bias are described.

Forest Inventory and Analysis

Each FIA unit selects a sample across the entire land base within its respective region. The sample is either random or systematic and is taken in at least two steps. The first step is to identify a large sample of points on aerial photographs. Most units stratify this sample into homogeneous groups. From the photo sample, a subsample is drawn for ground observation. FIA ground plots are made up of clusters of 3 to 10 subplots, centered on the photo point and spaced evenly over an acre.

All FIA units use plot clusters for ground observation, yet there are substantial inconsistencies in the way edge samples have been handled. Some of these sampling practices result in one or more of the following types of bias: underestimation of population totals, undersampling of conditions at the forest edge, and/or artificial inflation of "mixed" cover-type classifications. The seriousness of the bias is a function of edge frequency and is therefore unknown.

Until recently, most FIA units restricted field measurements to clusters where the center subplot was forested. No measurements were taken on clusters with nonforest centers even if some of the other subplots in the cluster overlapped into a forested condition. There are no inherent problems with this methodology in the single plot case, as long as the boundary bias was recognized and edge trees were properly weighted. However, in the cluster sampling case, estimation of the correct weights is extremely complex, so in many cases the bias was ignored. Thus the first source of bias resulted in underestimation of tree attribute totals, because some "forest" plots contained fragments of nonforest land and nothing was done to compensate.

To avoid underestimating tree attributes and to ensure an adequate sample of trees, several FIA units shifted plots away from the edge by moving the cluster center and/or systematically moving (or "substituting") nonforest subplots into forest (Figure 1). However, these protocols introduced a second source of bias by undersampling conditions at the forest edge. The magnitude of this bias depends on the population and the attributes observed. Any features or species that tend to occupy edge conditions will be underestimated. Number of trees and volume per unit-area are also likely to be underestimated (Gregoire and Scott 1990).

A third source of bias involves misclassification of stand attributes. Computer algorithms applied to the tree data collected on the cluster are used to classify stands into discrete categories such as cover type and size class. If cluster points

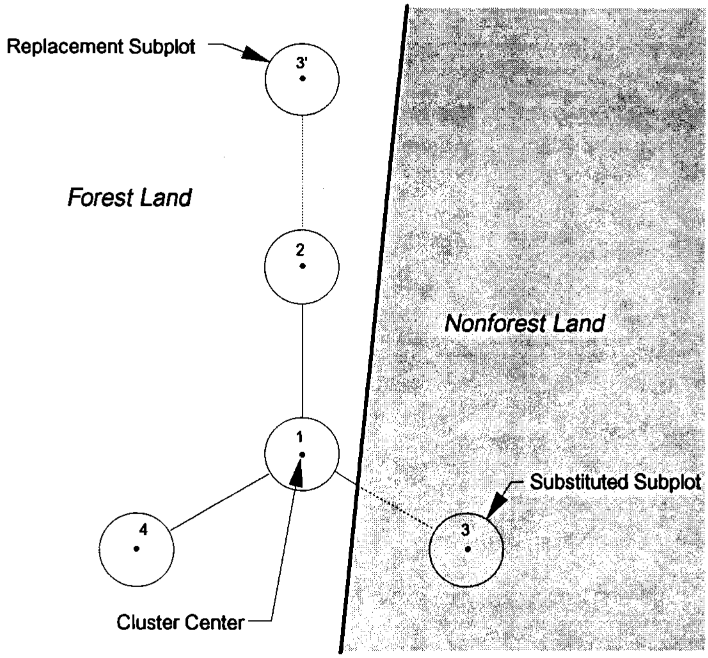

FIGURE 1. Example of a subplot falling on nonforest land that was replaced by a forested subplot.

were not moved but allowed to straddle cover-type boundaries, a "mixed" classification often resulted, even though two distinct types might be present. For example, a cluster crossing the boundary between a planted pine stand and a natural oak stand often resulted in a classification of mixed oak-pine. Thus, estimates of area by cover type were biased against pure types, artificially inflating estimates of mixed types. Some FIA units attempted to correct this problem by moving all subplots into the same forest condition, but at the expense of undersampling forest-forest edge.

FOREST HEALTH MONITORING

A national Forest Health Monitoring (FHM) Program has evolved in response to increasing concerns about the effect of various anthropogenic and natural stressors on U.S. forests (Palmer et al. 1991). FHM is jointly sponsored by the USDA Forest Service, U.S. Environmental Protection Agency, USDI Bureau of Land Management, Tennessee Valley Authority, USDA Natural Resources Conservation Service, many state agencies, and others. The primary function of FHM is to gather and maintain an objective data base capable of supporting appraisals of forest health at the regional and national levels.

Similar to FIA, the FHM program has established a systematic grid of sample locations across the United States. At each forested grid point, ground observations are recorded periodically from clusters of fixed-area subplots. Because both FIA and FHM faced the same boundary-bias issues, a meeting of federal, university, and corporate biometricians was convened in 1991 to address them. It was the consensus of the group that subplots should not be moved. This decision was the impetus for the boundary mapping techniques developed for the FHM program and recommended in this paper.

BOUNDARY MAPPING

Given the constraints that subplots never should be moved and that boundary-bias correction methods are impractical for plot clusters, boundary mapping was selected as the most appropriate way to attain the desired estimates. A variation of this procedure has been implemented successfully in Sweden since the early 1970s (Soderberg 1992). The mapping technique, which allows plots to cross multiple land use and cover types, is simple conceptually. Since plots are occupied and measured if *any* portion of the cluster is forested, without regard to the land use at plot center, estimates of tree characteristics such as volume require no correction for sampling near stand boundaries.

Boundary mapping also yields unbiased estimates of stand and area attributes, assuming that boundaries are mapped correctly. In contrast to the method of classifying whole clusters as a single entity, mapped plots are subdivided into various "condition classes" designed to yield more precise estimates of forest area. Combinations of variables used to define condition classes are made up of discrete variables which are commonly used to pre- or poststratify tree data (e.g., forest cover type, stand size, or physiography). Thus, the entire cluster is characterized (mapped) rather than just one or more points. This should result in more precise estimates of forest area and improved classification of conditions.

Boundary mapping is designed to compute the area of all conditions on each cluster so that estimates of forest area can be categorized by the variables of interest. Mapping of individual boundaries in the field is required only when the boundary between conditions crosses a subplot, and not when boundaries occur between subplots or when the entire cluster is in a single condition. To ensure that all tree data are compatible with stand-area data, it also is necessary to assign trees to their respective condition class.

FIELD PROCEDURES

The cluster design developed for FHM serves as a convenient demonstration of the recommended technique. Described briefly here, it is presented in greater detail by Scott (1993) and Bechtold et al. (1992). FHM clusters are composed of four points spaced 36.6 m (120.0 ft) apart (Figure 2). Each point serves as the center of a 0.0168 ha (1/24-ac) circular subplot for trees 12.7 cm (5.0 in.) dbh and larger. Each subplot also includes a 13.5 m^2 (1/300-ac) circular microplot for trees 2.54 to 12.7 cm (1.0 to 4.9 in.) dbh.

On the ground, clusters are established without regard to land use or forest cover. The condition class occupied by each subplot center is defined and recorded. Condition classes are assigned nominal labels (condition codes) that define combinations of five stand attributes: land use (forest vs. nonforest), forest type, stand origin (planted vs. natural), stand size (seedling/sapling, poletimber, sawtimber), and disturbance history. This list can be expanded to include variables such as ownership, physiography, or regional borders. When there is a change in

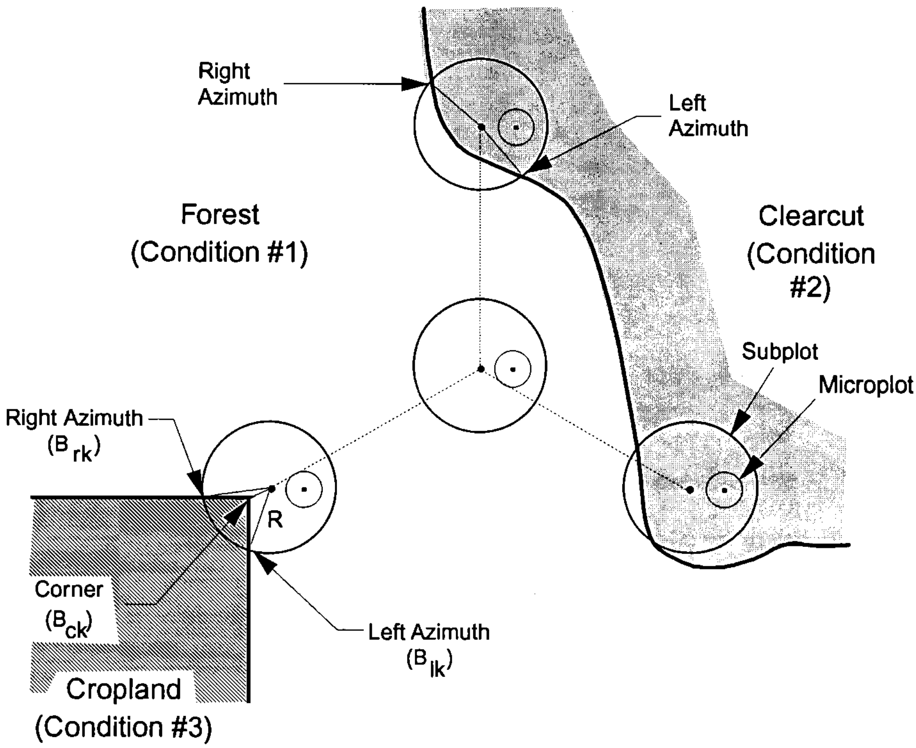

FIGURE 2. Example of boundary delineation on a FHM plot cluster.

one of these five elements, another condition class is defined. If there are two or more conditions on the same subplot and the boundary between them is *distinct*, the boundary is mapped relative to the subplot perimeter, and all trees tallied are assigned to their respective condition class. Individual microplots are mapped in a similar manner.

Subplot boundaries can be mapped in several ways. The method used by FHM is simple and usually requires just two azimuths—from the subplot center to the two points where the boundary intersects the subplot perimeter (Figure 2). The distances to the two perimeter points always are equal to the subplot radius. Also, one turning point can be used to designate a sharp bend or corner; this requires an additional azimuth and distance. Facing the boundary from the subplot center, the azimuths to the points where the boundary crosses the circumference are recorded as left and right azimuths. If applicable, a corner distance and azimuth are recorded as such. Multiple condition classes and boundaries are possible. FHM allows a maximum of nine condition classes per cluster and three boundaries per subplot. Other methods that map boundaries in greater detail are possible but are more complicated and difficult to remeasure consistently.

COMPUTING PLOT AREA BY CONDITION

To estimate the total area of a condition, it is necessary to calculate the proportion of each cluster located in that condition. If whole subplots are in or out of a condition, the estimates are simple. The cluster area, A_{ijk}, is 0 if condition k does not occur on subplot j of cluster i. If the subplot is entirely within condition k, then:

$$A_{ijk} = A_j = \pi R^2/43560 \text{ ac} \quad (1)$$

or

$$= \pi R^2/10000 \text{ ha}$$

where

R = radius of the subplot used to determine area by condition (in feet or meters), which is 24.0 ft for FHM

Computations are complicated only when a boundary crosses a subplot (Figure 2). The field crews record the following:

B_{lk} = azimuth (bearing) to the *left* intersection of the subplot circumference and the boundary for condition k (in clockwise degrees from North)

B_{rk} = azimuth to the *right* intersection of the subplot circumference and the boundary for condition k (in clockwise degrees from North)

B_{ck} = azimuth to the boundary *corner* for condition k (in clockwise degrees from North)

R_{ck} = distance (in feet or meters) to the *corner* of condition k

If the boundary has no corner, then the area of the fragment beyond the boundary line is:

$$A_{ijk} = 1/2\, R\, \{R\, Q_k - R \sin Q_k\}/43560 \text{ ac}$$
$$= R^2(Q_k - \sin Q_k)/87120 \text{ ac} \quad (2)$$

or

$$= R^2(Q_k - \sin Q_k)/20000 \text{ ha}$$

where

$$Q_k = \pi(B_{rk} - B_{lk})/180 \qquad \text{if } B_{rk} > B_{lk}$$
$$= 2\pi + \pi(B_{rk} - B_{lk})/180, \qquad \text{otherwise}$$

Letting $U = 87120$ for acres and $U = 20000$ for hectares, Equation (2) can be more generally written as:

$$A_{ijk} = R^2(Q_k - \sin Q_k)/U$$

Additional condition fragments are computed analogously. Fragment data are recorded so they do not partially overlap other fragments. In some cases, one fragment may be wholly within another (e.g., a narrow road divides a subplot between the subplot center and the subplot perimeter). In that case, the area of the larger fragment must be reduced by the area of the smaller fragment. The area of the remaining portion of the subplot (encompassing subplot center) is determined by subtraction from A_j, Equation (1).

In the case where there is a corner (turn) in the boundary, the following formula is used:

$$A_{ijk} = R\{R\, Q_k - R_{ck}(\sin Q_{rk} + \sin Q_{lk})\}/U \quad (3)$$

where

$$Q_{rk} = \pi(B_{rk} - B_{ck})/180$$

$$Q_{lk} = \pi(B_{ck} - B_{lk})/180$$

This equation applied to boundaries with and without corners. If the boundary does not have a corner point recorded, set $R_{ck} = R$ and $B_{ck} = B_{lk}$. Thus, Equation (2) is a special case of (3).

Equation (3) applies when the angle formed by the three boundary points is less than 180° (π radians). If not, then the subplot center falls within the pie-shaped wedge formed by the boundary (Figure 3). The roles of the left and right corners are reversed to compute the pie-shaped complementary fragment, which is then subtracted from the whole. Thus, the area beyond the boundary is estimated as:

$$A_{ijk} = A_j - R\{R\, Q_k - R_{ck}(\sin Q_{rk} + \sin Q_{lk})\}/U \qquad (4)$$

where

$$Q_{rk} = \pi(B_{lk} - B_{ck})/180$$

$$Q_{lk} = \pi(B_{ck} - B_{rk})/180$$

With these formulas, virtually any combination of fragments formed by condition boundaries can be accommodated.

Inconsistencies in Boundary Data

Field crews assign each tree to a condition and a condition to each fragment. It is desirable to record horizontal distance and azimuth to the center of each tree to verify that the tree location and condition observations are consistent with the fragment location and condition observations. Similarly, it is possible to compare boundary points between fragments to check for overlapping fragments.

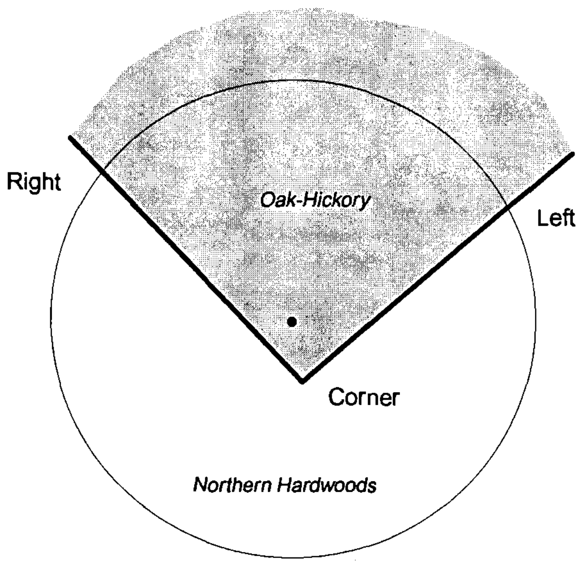

Figure 3. Example of a boundary whose angle is greater than 180°.

Errors in Locating Trees Within Fragments

With added complexity comes the increased opportunity for error. However, the boundary mapping method also affords the opportunity to build in checks to detect potential errors. This section describes a method which simultaneously checks for correct boundary location data, tree location data, and assignment of the tree to a fragment. It checks to make sure that the tree location corresponds to a fragment of the same condition class.

The approach here is to shift the origin from the subplot center to either the corner point or, in the single-boundary case, to one created just beyond the right point. This is done to make the tests easier to perform. The shift is accomplished as follows:

Single Boundary	*Double Boundary*	
$CX = R\cos(B_{rk})$	$CX = R_{ck}\cos(B_{ck})$	(5)
$CY = R\sin(B_{rk})$	$CY = R_{ck}\sin(B_{ck})$	(6)
$RR = R + 1$	$RR = R$	

The (x,y) coordinates of the corner point are given by (CX, CY). In the single-boundary case, the right point is shifted out by 1 ($RR = R + 1$), and the corner point assumes the old position of the right point. This allows us to treat the single-boundary case in the same way as the double-boundary case for the remaining steps.

Given the distance, D_t, and the azimuth, B_{rk}, to a tree, the (x,y) coordinates of the left boundary point, (X_l, Y_l), the right boundary point, (X_r, Y_r) and the tree, (X_t, Y_t) are recomputed using the corner point as the origin (Figure 4).

$$X_l = R\cos(B_{lk}) - CX \quad (7)$$

$$Y_l = R\sin(B_{lk}) - CY \quad (8)$$

$$X_r = RR\cos(B_{rk}) - CX \quad (9)$$

$$Y_r = RR\sin(B_{rk}) - CY \quad (10)$$

$$X_t = D_t\cos(B_t) - CX \quad (11)$$

$$Y_t = D_t\sin(B_t) - CY \quad (12)$$

The next step is to compute the azimuth from the new origin to the left and right points, and to the tree ($l = l, r, t,$ respectively):

$$\begin{aligned} B_l &= \tan^{-1}(Y_l/X_l) && \text{if } X_l > 0 \text{ and } Y_l > 0 \\ &= \pi + \tan^{-1}(Y_l/X_l), && \text{if } X_l < 0 \\ &= 2\pi + \tan^{-1}(Y_l/X_l), && \text{if } X_l > 0 \text{ and } Y_l < 0 \end{aligned} \quad (13)$$

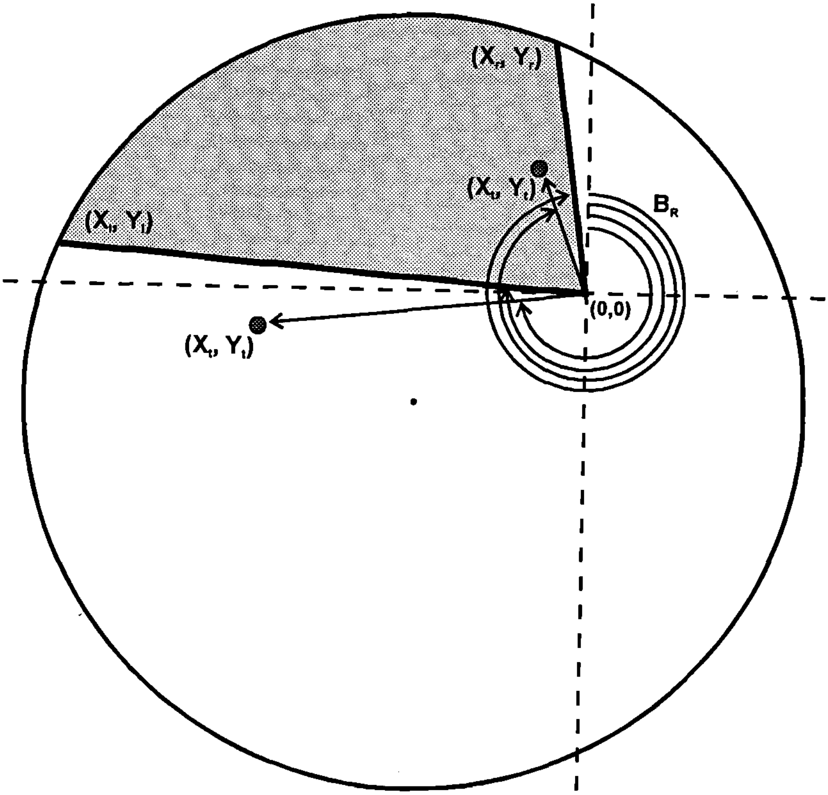

FIGURE 4. Example of determining which trees are in a fragment.

Finally, verify that the azimuth to the tree is between the azimuths to the left and right points:

If $B_r \geqslant B_l$, *then:*

If $B_l \leqslant B_t \leqslant B_r$, then the tree is in the fragment

If $B_r < B_l$, *then:*

If $B_l < B_t$ or $B_t < B_r$, then the tree is in the fragment

If the tree is in the fragment but the condition code of the fragment and the tree are not equal, check the observations for inconsistencies. If the tree is not in any fragment of the same condition code, compare it to the condition code for the subplot center. If they are not equal, then the observations must be checked.

When identifying inconsistencies, it is useful to indicate how far out the tree is. Because actual boundaries rarely are straight, some flexibility in tree location must be allowed. The relevant distances are those extending from the tree to the boundaries and to the corner point, if any. The distance from the tree to the corner point can be computed as:

$$D_t' = (X_t^2 + Y_t^2)^{1/2} \qquad (15)$$

The perpendicular distance from the tree to the lines through the new origin and each of the two boundary points are computed as:

$$OUT_l = |D_t' \sin(B_t - B_l)| \tag{16}$$

$$OUT_r = |D_t' \sin(B_t - B_r)| \tag{17}$$

Note that this distance is from a straight-line extension of the boundary so that the distance can be computed as a perpendicular distance. The closest distance to the actual boundary may be to the corner, D_t'. This method applies to any of the observations that must be checked, (X_t, Y_t), not just trees. Thus, it can be used to determine whether the corner point of another fragment is located within the current fragment.

Errors in Recording Boundary Data

A common field error is transposition of the left and right azimuths. If this happens, the fragment will include the subplot center—a situation that by definition cannot happen (the area of the fragment containing the subplot center is not derived explicitly but computed by subtracting the area of the fragments from the total). In the case where there is no corner point, the azimuths have been reversed if Q_k is greater than π (180°). In the case where the boundary does have a corner, then:

If $Q_k > \pi$ then: (18)

If $\sin(B_{lk} - B_{ck}) > 0$ or $\sin(B_{ck} - B_{rk}) > 0$ then "Values Switched"

If $Q_k \leq \pi$ then:

If $\sin(B_{rk} - B_{ck}) < 0$ or $\sin(B_{ck} - B_{lk}) < 0$ then "Values Switched"

If values are switched, then interchange the values of B_{lk} and B_{rk} and recompute.

Other checks for inconsistencies involve comparisons of conditions assigned to the fragments. If a fragment is not wholly contained within another fragment, it should not have the same condition as the subplot center. A fragment within another fragment should not have the same condition as the fragment it overlaps. Condition fragments should reside either wholly inside or outside of other fragments. This is checked by determining whether all or none of the two or three boundary points are located within another fragment. The location of the left and right boundary points and of the corner point is substituted into the tree location formulas (5–14) to perform this comparison.

ESTIMATION

In most regional forest survey applications, the values of primary interest are total area in the region by condition, and the totals of various tree attributes by condition or across the region. Once these values are determined, the totals can be placed on a per-unit-area basis. Typically, the attribute totals are divided by the total forest area, thus forming ratio estimators. Estimated attribute totals by condition also can be divided by the estimated total area in the condition. One application of these ratio estimates by condition is in conjunction with mapped polygons in the same condition class. Estimates of cluster-level ratios can be generated for modeling purposes, such as developing relationships between understory species composition to overstory basal area by condition.

Population Estimators of Area

Using the areas by condition developed in Equations (1) to (4), the proportion of cluster i falling in condition k is computed as the total sample area within condition k in cluster i divided by the total sample area:

$$P_{ik} = \frac{\sum_{j}^{m} A_{ijk}}{\sum_{j}^{m} A_j} \tag{19}$$

where

A_{ijk} = area of subplot j in condition k on cluster i

A_j = total area of subplot j which is a constant for all clusters [see Equation (1)]

m = number of subplots in a cluster

This value is then expanded to the population level based on the design. If simple random sampling of clusters is used, the estimated total area in condition k is the average of the cluster proportions times the total population area:

$$\hat{A}_k = A_T \sum_{i}^{n} P_{ik}/n \tag{20}$$

where

A_T = total land area in the population (region)

n = number of clusters sampled

If other sampling designs are used, such as double sampling for stratification (Cochran 1977, Eq. 12.32) or unequal probability sampling (Raj 1968, Eq. 3.22), then equations similar to (20) can be applied. The methods used in FHM are described in Palmer et al. (1991).

The variance estimator also is design dependent but needs only to consider the variance between clusters and not the variance within clusters. This results from the fact that the sample-based variance estimator of the between-cluster variation includes within-cluster variation and that in extensive forest inventories the sampling fractions are small (Cochran 1977, Eq. 10.23). Thus, in the simple random-sampling case, the variance of the estimated total area in condition k is:

$$v(\hat{A}_k) = A_T^2 \frac{\sum_{i}^{n} (P_{ik} - \hat{A}_k/A_T)^2}{n(n-1)} \tag{21}$$

Population Estimators of Tree Attributes

Population estimators of tree attribute totals by condition class are developed similarly. First, the cluster value for the attribute of interest in the condition of interest is expressed on a per-unit-area basis (Y_{ik}).

$$Y_{ik} = \frac{1}{m}\sum_{j}^{m} Y_{ijk} = \frac{1}{m}\sum_{j}^{m}\sum_{t}\frac{Y_{ijkt}}{A_{ijt}} \tag{22}$$

where

Y_{ijk} = attribute of subplot j in condition k on cluster i

Y_{ijkt} = attribute of tree t in condition k on subplot j in cluster i

A_{ijt} = plot size in acres (ha) for tree t on subplot j in cluster i (constant for fixed-area subplots)

The attributes of interest can be any measured tree characteristic (e.g., volume, basal area, or numbers of cavities). The condition of interest is any combination of condition-class variables recorded on the plots (e.g., undisturbed natural pine stands). If the attribute/condition of interest does not qualify, the observation is set to 0. The plot size is design dependent, such as in variable-radius (Bitterlich) sampling versus fixed-area sampling with concentric plots. This estimator, (22), also can be rewritten more generally to expand trees by the total sample area within the cluster from which trees of similar characteristics could have been sampled, thus including the case of overlapping tree circles (Van Deusen and Grender 1989). Note that Y_{ik} is on a per-unit-area sampled basis and not on a per-unit-area of condition basis. The latter estimator is given in (26).

Assuming simple random sampling, the population total and its variance are estimated by substituting Y_{ik} for P_{ik} in Equations (20) and (21), respectively.

Per-Unit-Area Estimators

To express an attribute on a per-unit-area basis, the estimated population total of the attribute is divided by the estimated total area of interest. For example, the average number of merchantable maple stems per hectare of northern hardwood stands is estimated as the total number of merchantable maple stems in northern hardwood conditions divided by the total area of northern hardwood conditions. This is a ratio-of-means estimator with the variance given by Cochran (1977, Chap. 6). The estimator usually has a small bias, and its variance is a good approximation.

$$\hat{R}_k = \hat{Y}_k/\hat{A}_k = \bar{Y}_k/\bar{A}_k = \frac{\sum_i Y_{ik}/n}{\sum_i A_{ik}/n} \tag{23}$$

A variance estimator from Cochran (1977, Eq. 6.13) is:

$$v(\hat{R}_k) = (v(\hat{Y}_k) + \hat{R}_k^2\, v(\hat{A}_k) - 2\,\hat{R}_k \operatorname{cov}(\hat{Y}_k,\hat{A}_k))/\hat{A}_k^2 \tag{24}$$

Assuming simple random placement of clusters, the covariance is:

$$\operatorname{cov}(\hat{Y}_k,\hat{A}_k) = A_T^2 \frac{\sum_i^n (Y_{ik} - \bar{Y}_k)(A_{ik} - \bar{A}_k)}{n(n-1)} \tag{25}$$

Submerchantable stems are measured on a microplot, where boundary data are recorded with the same mapping protocol described for subplots. Several approaches are possible when developing population-level ratio estimators for combined data obtained from both subplot and microplots. Ratios from the subplots and microplots can be computed separately, then summed. If this method is used, the area estimate in the denominator of the microplot ratio can use area estimates obtained from either the subplots or microplots. However, this produces the sum of the two correlated ratio estimators, the variance of which is likely to be intractable. A better approach is to estimate the total number of stems in the population and divide by the area as estimated from the subplot areas; then use Equations (23) and (24) as before.

When modeling is of interest, per-unit-area basis values are required for each cluster. These are simply computed as:

$$R_{ik} = Y_{ik}/P_{ik} \tag{26}$$

Note that the mean of these ratios will not yield the same estimate as the ratio-of-means estimator, (23).

DISCUSSION

The advantages and disadvantages of boundary mapping compared to other proposed solutions to edge-bias problems are discussed in detail in Hahn et al. (1995). The following details require further elaboration.

If a particular condition is sampled by only a small fragment of a cluster, little sample-based information is available to classify the condition. This situation will occur frequently when boundaries cross a subplot, especially small subplots. In such cases, field crews might be permitted to make subjective classifications at the expense of introducing potential measurement error. If this is not acceptable, the subplot should be enlarged by some predetermined factor to generate additional data which are used for classification purposes only (Soderberg 1992).

In situations where the use of prism sampling is advantageous, it is possible to map plots by imposing an upper limit at which the radius of the variable plot is fixed (Hahn et al. 1995). This avoids the selection of trees in conditions for which no area is observed and fixes the area that must be mapped. However, it should be noted that misclassification is likely when tree data from mapped prism plots are used to compute area attributes by condition class (e.g., stocking, stand size, forest type). On plots fragmented into multiple conditions, prism sampling leads to a situation where it is not possible to sample the full range of tree sizes in all conditions. The range of excluded sizes is a function of the prism factor and the distance of the plot fragment from point center. The only solution is to engage some form of supplementary sampling for classification purposes on plots with boundaries. The extra effort may nullify the advantages of prism sampling.

Analytical processes are more complicated with mapped clusters, unless the entire cluster is in a single condition. When attributes are estimated by condition, those portions of the cluster not in the condition are set to zero or "missing." As a result, more "bookkeeping" is required. With subsampling for other resource attributes (e.g., regeneration data from microplots), some conditions may have no subsample data. These problems are negligible for regional aggregations of the data but could be important for modeling relationships between observations within the cluster.

Most FIA units stratify ground samples based on aerial-photo stratification. In applications where stratification by land use and forest cover class is used to reduce the variance of forest area and volume estimates, the use of mapped clusters becomes more complicated. Whole clusters could be assigned to strata based on plurality of the cluster occupied by a given condition class. Of course, the elements or criteria that comprise a condition class must be compatible with the elements that define the strata. This method has the advantage of using the current strata but increases the variability within the strata. Another approach is to create one or more classes for clusters that straddle conditions (strata), thereby isolating the heterogeneity into a "mixed" stratum. Alternatively, if subplots within the cluster are spaced widely enough, each point in the cluster can be classified individually.

Problems with sampling at the edge are not entirely eliminated by the methodology described here. Plots on population boundaries between regional survey units such as state, county, or regional borders require special attention. There are several alternatives, none of which is entirely satisfactory. The simplest is to ignore the population boundary and sample the cluster as usual. In many cases, the population boundary is not apparent on the ground anyway. This approach works well as long as the adjacent population does not differ significantly from the sample population. Alternatively, the portions of the cluster out of the population can be treated as null. This avoids including something from the adjacent population that does not occur in the sample population, but results in an underestimate of the population totals. With small sampling fractions and small clusters, the probability of having a cluster cross into the adjacent population is minimal.

When the sampling grid is relatively sparse, such as with FHM, estimates of forest area obtained from more intensive observations from photographs or satellite imagery may be more reliable. It may be desirable to constrain population estimates of area and tree attributes based on other independent estimates of forest area. It also may be desirable to use other estimates to avoid the possibility of different agencies reporting different estimates. Population totals can be conditioned by a ratio estimator, as shown by the following example:

$$\hat{A}'_k = \hat{A}_k(\hat{A}_I/\hat{A}_F) = \hat{A}_I\,(\hat{A}_k/\hat{A}_F) = \hat{A}_I\,\hat{R}_k \tag{27}$$

where

$\hat{A}_I$ = independent estimate of total forest area

$\hat{A}_F$ = estimate of total forest area from sample

The variance is found in two steps. The first step is to estimate the variance of the ratio estimator, which is given in (24). The second step is to approximate the variance of the product of two independent estimators as:

$$v(\hat{A}'_k) \doteq \hat{A}_I^2\, v(\hat{R}_k) + \hat{R}_k^2\, v(\hat{A}_I) \tag{28}$$

Another method of constraining the estimators is presented in Li et al. (1990). Unlike the case treated there, the estimator used here, (20), is additive, i.e., the sum of all forested conditions equals the estimated total forest area.

A benefit of the mapped cluster method is the representative sample of edge conditions. This is important for assessments of wildlife habitat and forest succession. As opposed to moving clusters and subplots to keep the sample within, this method provides information on the juxtaposition of stands and other land uses. If plots are remeasured, it provides methods for estimating the area by type of change, as well as changes in forest attributes. Because each cluster can provide information on a number of conditions, the precision of area by condition class is improved while virtually all of the bias is eliminated. Also, the improvement in precision is likely to extend to many of the other resource attributes.

RECOMMENDATIONS

For large-scale forest surveys, plot clusters should be installed in a fixed pattern, with subsequent sampling of any forest conditions located on individual subplots, regardless of conditions at the cluster center. Condition boundaries that cross subplots should be located spatially (mapped), with assignment of individual trees to their respective conditions.

In situations where the use of variable-radius subplots is advantageous, we suggest imposing an upper limit at which the radius of the variable plot is fixed (Hahn et al. 1995). This avoids selecting trees in conditions for which no area is observed, fixes the area that must be mapped, and reduces the incidence of missing large, distant trees. However, in some cases, use of variable plots complicates the computation of area attributes that are based on classifications of tree data.

The mapping techniques described in this paper eliminate the inherent biases associated with two practices commonly used in regional surveys—moving subplots away from edges and averaging data across multiple conditions. As with any survey, correct field application is essential. With the mapped design, proper location of boundaries is crucial to accurate estimates of area by condition. Since the decision to place a boundary is based on the same criteria as the decision to move a subplot, a switch from subplot substitution to boundary mapping is not likely to result in any more boundary identification and measurement error than existed under substitution protocols. With other boundary-bias correction methods, errors in boundary location can affect the condition class assignment, the weight assigned to the tree, and the estimate of the population total. However, with the mapped technique, errors in boundary location can affect the condition class assignment, but not the individual tree weight nor the population total.

LITERATURE CITED

BARRETT, J.R. 1965. Corrections for edge effect bias in point sampling. For. Sci. 10:52–55.

BECHTOLD, W.A., V.J. LABAU, W. MCLAIN, and P. ROGERS. 1992. Site classification, growth, and regeneration. *In* Forest health monitoring field methods guide (national guide). Conkling, B.L., and G.E. Byers (eds.). U.S. Environ. Prot. Agency, Off. of Res. and Dev. (internal report). Forestry Sciences Lab, P.O. Box 12254, Research Triangle Park, NC 27709.

COCHRAN, W.G. 1977. Sampling techniques. Ed. 3. Wiley, New York. 428 p.

FINNEY, D.J., and H. PALCA. 1948. The elimination of bias due to edge-effects in forest sampling. Forestry 23:31–47.

FOWLER, G.W., and L.G. ARVANITIS. 1981. Aspects of statistical bias due to the forest edge: Horizontal point sampling. Can. J. For. Res. 11:334–341.

GREGOIRE, T.G., and C.T. SCOTT. 1990. Sampling at the stand boundary: A comparison of the statistical performance among eight methods. P. 78–85 *in* Research in forest inventory, monitoring, growth and yield. IUFRO World Congr. Virginia Polytech. and State Univ., Publ. FWS-3-90.

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

HAHN, J.T., C.D. MACLEAN, S.L. ARNER, and W.A. BECHTOLD. Procedures to handle inventory cluster plots that straddle two or more conditions. For. Sci. Monogr. 31:12–25.

LI, H.G., H.T. SCHREUDER, and C.T. SCOTT. 1990. Combining estimates that are both in error subject to marginal constraints. Can. J. For. Res. 20:1675–1679.

PALMER, C.J., et al. 1991. Monitoring and research strategy for forests—Environmental Monitoring and Assessment Program. U.S. Environ. Prot. Agency. Publ. EPA 600/4-91/012. 10 p.

RAJ, D. 1968. Sampling theory. McGraw-Hill, New York. 302 p.

SCHMID-HAAS, P. 1969. Stichproben am waldrand (Sampling at the edge of the forest). Eidgen. Anst. Forstl. Versuchswes. Mitt. 45:234–303.

SCOTT, C.T. 1993. Optimal design of a plot cluster for monitoring. P. 233–242 *in* The optimal design of forest experiments and forest surveys, Rennolls, K., and G. Gertner (eds.). London, UK., Univ. of Greenwich, Sch. of Math., Stat. and Computing.

SCOTT, C.T., A.R. EK, and T.R. ZEISLER. 1983. Optimal spacing of plots comprising clusters in extensive forest inventories. P. 707–710 *in* Renewable resource inventories for monitoring changes and trends. Soc. Am. For. Publ. SAF 83–14.

SCOTT, C.T., D.L. CASSELL, and J.W. HAZARD. 1993. Sampling design of the U.S. National Forest Health Monitoring Program. P. 150–157 *in* Ilvessalo symposium on national forest inventories. Nyyssonen, A., S. Poso, and J. Rautala (eds.). Finnish For. Res. Inst. Res. Pap. 444.

SODERBERG, U. 1992. Pers. comm. Swedish Univ. of Agric. Sci., Dep. of For. Surv. Umea, Sweden.

USDA Forest Service. 1987. Forest Service resource inventory: An overview. USDA For. Serv., For. Inventory and Econ. Res., Washington, DC. 29 p.

VAN DEUSEN, P.C., and J.M. GRENDER. 1989. An alternative estimator for horizontal point clusters. For. Sci. 35:257–262.

Manuscript received October 6, 1993

AUTHORS AND ACKNOWLEDGMENTS

Charles T. Scott is Research Forester, Northeastern Forest Experiment Station, Delaware, OH 43015; William A. Bechtold is Research Forester, Southern Research Station, Asheville, NC 28802. The authors acknowledge the ideas of the Forest Inventory and Analysis Committee on plot design, of others involved in the review of the Forest Health Monitoring design, and the critical comments of Hans T. Schreuder of the Rocky Mountain Forest and Range Experiment Station.